73147

From:	Li, Ruixiang
Sent:	Tuesday, August 13, 2002 11:13 AM
To:	STIC-Biotech/ChemLib
Subject:	Sequence search of Application NO:09/898,586

Please do a standard search on SEQ ID NO: 24 against both commercial and interference nucleic acid databases.

Thank you very much!

Ruixiang Li
GAU 1646
CM1 10E18
Mail Box 10C01
306-0282

(1)

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Searcher: _____	NA Sequences: _____	STN: _____
Phone: _____	AA Sequences: _____	DIALOG: _____
Location: _____	Structures: _____	Questel/Orbit: _____
Date Picked Up: _____	Bibliographic: _____	DRLink: _____
Date Completed: 8/23/02	Litigation: _____	Lexis/Nexis: _____
Searcher Prep/Review: _____	Full text: _____	Sequence Sys.: _____
Clerical: _____	Patent Family: _____	WWW/Internet: _____
Online time: _____	Other: _____	Other (specify): _____

```
OM of: US-09-898-586-24 to: GenEmbl:*    out_format : pfs

Date: Aug 21, 2002   8:18 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL-frame+_p2n.model -DEV-xlh
 -Q-/cgn2_1/USPTO_spool/US09898586/runat_16082002_205735_20549/app_query.fasta_1.373
 -DB-GenEmbl -QFMT-fastap -SUFFIX-rge -GAPOP-12.000 -GAPEXT-4.000
 -MINMATCH-0.100 -LOOPCL-0.000 -QGAPOP-4.500
 -MINMATCH-0.100 -LOOPEXT-0.000 -QGAPOP-4.500
 -QGAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500 -FGAPOP-6.000
 -FGAPEXT-7.000 -YGAPOP-10.000 -YGAPEXT-0.500 -DELOP-6.000
 -DELEXT-7.000 -START-1 -MATRIX-blosum62 -TRANS=human40.cdi
 -LIST-45 -DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN=0
 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs -NORM-ext -HEAPSIZE-500
 -MINLEN-0 -MAXLEN-2000000000 -USER-US09898586_@CGN1_1_4134
 -NCPU-6 -ICPU-3 -LONGLOG -DEV_TIMEOUT-120 -WARN_TIMEOUT-30
 -NO_XLPXY -WAIT -THREADS-1

Search information block:
Query: US-09-898-586-24
Query length: 310
Database: GenEmbl:*
Database sequences: 1797656
Database length: 18733333701
Search time (sec): 1826.930000
```

score_list:

Sequence	Strd	Orig	zScore	EScore	Len	Documentation
gb_pat:AX078373	+	1609.00	2439.04	1.7e-127	974	! AX078373 Sequence 41 from Paten
gb_pat:AX197456	+	1609.00	2438.73	1.8e-127	1012	! AX197456 Sequence 21 from Pate
gb_pri:AC004889	-	1609.00	2398.40	3.1e-125	130030	! AC004889 Homo sapiens clone
gb_pat:AX242191	+	1602.00	2398.79	6.4e-127	939	! AX242191 Sequence 939 from Pate
gb_pat:AX241442	+	1584.00	2401.45	2.1e-125	930	! AX241442 Sequence 190 from Pate
gb_pat:AX241473	+	1584.00	2401.45	2.1e-125	930	! AX241473 Sequence 221 from Pate
gb_pat:AX197438	+	1584.00	2400.52	2.4e-125	1040	! AX197438 Sequence 34 from Pate
gb_pat:AX197469	+	1584.00	2400.52	2.4e-125	1040	! AX197469 Sequence 34 from Pate
gb_pat:AX135904	-	1584.00	2366.96	1.5e-123	140915	! AX135904 Human DNA sequence f
gb_pri:AL135904	-	1584.00	2359.75	4.5e-123	140915	! AL135904 Human DNA sequence f
gb_pri:AC005587	+	1584.00	2359.75	4.5e-123	1014	! AC005587 Homo sapiens PAC cl
gb_pri:AX197458	+	1568.00	2376.43	5.3e-124	1014	! AX197458 Sequence 23 from Pate
gb_pat:AX074444	+	1330.00	2015.57	6.6e-104	933	! AY073440 Homo sapiens olfactory
gb_pat:AC091768	+	1265.00	1877.60	3.2e-96	104955	! AC091768 Homo sapiens BAC cl
gb_ro:AY073100	+	1265.00	1887.96	8.5e-97	933	! AY073100 Mus musculus olfactory
gb_pri:AC074386	+	1138.00	1688.45	1.1e-85	66610	! AC074386 Homo sapiens BAC clo
gb_pri:AF399598	+	1129.00	1713.23	4.6e-87	650	! AF399598 Homo sapiens clone OR2
gb_pat:AX242188	+	1126.00	1705.70	1.2e-86	930	! AX242188 Sequence 936 from Pate
gb_pat:AX242190	+	1124.00	1702.11	1.9e-86	993	! AX242190 Sequence 938 from Pate
gb_pat:AX078369	+	1123.00	1695.69	4.3e-86	1792	! AX078369 Sequence 37 from Pate
gb_pat:AX073723	+	1067.00	1615.75	1.2e-80	966	! AY073723 Mus musculus olfactory
gb_ro:AY092423	+	1048.50	1582.89	8.3e-80	1713	! AR092423 Sequence 1 from patent
gb_ro:AY073724	+	1046.00	1583.85	7.4e-80	966	! AY073724 Mus musculus olfactory
gb_ro:AY073443	+	1040.00	1574.95	2.3e-79	942	! AY073443 Mus musculus olfactory
gb_pat:AX241995	+	1002.00	1502.03	2.7e-75	942	! AX241995 Sequence 743 from Pate
gb_ro:AY073820	+	1002.00	1517.30	3.8e-76	933	! AY073820 Mus musculus olfactory
gb_htg:AC076959	+	991.00	1457.74	7.8e-73	162320	! AC076959 Homo sapiens chromo
gb_htg:AC073798	+	991.00	1499.12	3.9e-75	927	! AY073798 Mus musculus olfactory
gb_ro:AY073440	+	990.00	1476.28	2.2e-74	988	! AY073440 Mus musculus olfactory
gb_ro:AY073441	+	974.00	1474.76	8.8e-74	933	! AY073441 Mus musculus olfactory
gb_ro:AY073442	+	974.00	1473.24	1.1e-73	933	! AY073442 Mus musculus olfactory
gb_pat:MUSODORECA	+	965.00	1463.88	3.5e-73	667	! L14566 Mouse (K7) odorant recep
gb_htg:AC076959	-	947.50	1391.66	3.7e-69	162320	! AC076959 Homo sapiens chromo
gb_pat:AX241886	+	946.00	1431.75	2.2e-71	988	! AX241886 Sequence 634 from Pate
gb_ro:AY073797	+	943.00	1427.67	3.7e-71	933	! AY073797 Mus musculus olfactory
gb_pat:AX242170	+	920.00	1392.70	3.3e-69	937	! AX242170 Sequence 918 from Pate
gb_pri:AC004889	+	920.00	1351.73	6.3e-67	130030	! AC004889 Homo sapiens clone
gb_pat:AX241944	+	881.00	1333.38	6.6e-66	945	! AX241944 Sequence 692 from Pate
gb_pat:AX242035	+	854.00	1296.48	7.5e-64	576	! AX242035 Sequence 783 from Pate
gb_pat:AX241845	+	852.50	1220.04	1.7e-59	950	! AX241845 Sequence 593 from Pate
gb_ro:AY073997	+	851.00	1288.10	2.2e-63	912	! AY073997 Mus musculus olfactory
gb_htg:AC090838	-	840.00	1250.72	2.6e-61	10985	! AC090838 Homo sapiens chr
gb_htg:AC106354	-	838.50	1226.82	5.7e-60	148407	! AC106354 Rattus norvegic
gb_htg:AC106354	+	832.00	1216.95	2.0e-59	148407	! AC106354 Rattus norvegic
gb_pat:AX241887	+	813.00	1230.12	3.7e-60	941	! AX241887 Sequence 635 from
gb_htg:AC068994_3	-	796.50	1166.97	1.2e-56	92197	! Continuation (4 of 4) of

seq_name: gb_pat:AX078373

seq_documentation_block:

```
LOCUS       AX078373         974 bp    DNA     linear   PAT 22-FEB-2001
DEFINITION  Sequence 41 from Patent WO0107612.
ACCESSION   AX078373
VERSION     AX078373.1  GI:13158042
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 974)
AUTHORS     Au-Young,J., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
            Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
TITLE       Receptors and associated proteins
JOURNAL     Patent: WO 0107612-A 41 01-FEB-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..974
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID NO: 22792267CB1"
BASE COUNT      206 a    270 c    214 g    284 t
ORIGIN
```

alignment_scores:
```
Quality: 1609.00                         Length:      310
Ratio:     5.190                         Gaps:          0
Percent Similarity: 100.000   Percent Identity: 100.000
```

alignment_block:
```
1. .974
```

US-09-898-586-24 x AX078373

Align seg 1/1 to: AX078373   from: 1   to: 974

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr    17
    |||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCC  50

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
    |||||||||||||||||||||||||||||||||||||||||||||||||
 51 CGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCT 100

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuLeuIleSerLeu    50
    |||||||||||||||||||||||||||||||||||||||||||||
101 ACGTCTTCACCCTGCTGGGAAACGGGACCATACTGGGCTCATCTCACTG  150

 51 AspSerArgLeuHisSerAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
151 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGCCGGT  200

 67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 CGTCGACATGGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACC  250

 84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr    100
    |||||||||||||||||||||||||||||||||||||||||||||||
251 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGCAGACC  300

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe  117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 TTTCTGTTTTCCACTTTCGCTGTCACAGAATGTCTCCTCCTGGTGGTGAT 350

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
```

seq_name: gb_pat:AX197456

seq_documentation_block:

```
LOCUS       AX197456                1012 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 21 from Patent WO0151632.
ACCESSION   AX197456
VERSION     AX197456.1  GI:15387837
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1012)
AUTHORS     Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,
            Spytek,K.A. and Li,L.
TITLE       Odorant receptor polypeptides and nucleic acids encoding same
JOURNAL     Patent: WO 0151632-A 21 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1..1012
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      219 a    281 c    215 g    297 t
ORIGIN
```

alignment_scores:

```
                         Quality:  1609.00             Length:  310
                         Ratio:       5.190            Gaps:      0
            Percent Similarity: 100.000     Percent Identity: 100.000

alignment_block:
US-09-898-586-24 x AX197456

US-09-898-586-24:   ..
Align seq 1/1  to: AX197456      from: 1   to: 1012
```

```
351 GTCCTATGATCTGTACGGTGGCCANCTGCCACCCCTCCGATATTTGGCCA 400
    |||||||||||||||||||||||||||||||||||||||||||||||||||
134 leMetThrThrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150

401 TCATGACCTGGAGAGTCTGCATCCATCACCCCTCCGCGGTGACTTCCTGT 450

151 GlyValLeuLeuSerLeutIleHisLeuValLeutLeuLeuProLeuProPh 167
451 GGAGTCCTTTTATCCTTGANTCANCTTGTGTTACTTCTACCTTTACCTT 500

167 eCysArgProGlnLysIleThrYrHisPhePheCysGluIleLeuAlaVall 184
501 CTGTAGGCCCCAGAAAAATTATCACTTTTTTTGTGAAATCTTGGCTGTTC 550

184 eulyLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
551 TCAAACTTGCCTGTGCAGATACCCACATCAAATGAGAACATGGTCTTTGGGC 600

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
601 GGAGCAAATTGCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATA 650

217 rMetCysIleLeuCysAlaIleIleLeuGlnIleGlnInSerArgGlnValGlnA 234
651 TATGTGCATCCTCTGTGCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGA 700

234 rglyLysAlaPheCysThrCysPheSerHisLeuCyssValLeuGlyLeuPhe 250
701 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTGTGTGATGGAACCCA 750

251 TyrGlyThrAlalleIleMettYrvalGlyProArgTyrGlyAsnProly 267
751 TATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCATG 800

267 sGluGlnlyLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMett 284
801 GGAGCAAGAGAAATTCTCCTGCTGTTTCACAGCCTGTTTAATCCCATGC 850

284 euAsnProLeulleCysSerLeuArgAsnSerGluValLysAsnAlaIleu 300
851 TCAACCCCTTATCTGTAGTCTTAGGAACTCAGAAGGTGAAGAATACTTTG 900

301 LysArgValLeuGlyValGlyValGluArggAlaLeu 310
901 AAGAGAGTGCTGGGAGTAGGAAAGGGCTTTA 930
```

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGGGAGACAATATAACATCCATCACAGAGTTCCTTCTACTGGGATTTCC 54

 54 oValGlyProArgglleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34
103 CGTTGGGCCCAAGGAATTCAGATGCTGCTCTTTGGGCTTTTCCCGTTTCTT 153

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyIleIleSerLeu 50
154 ACGTCTTCACCCTCCTGGGAAACGGGACCAATACTGGGTCATCTCACTG 203

 51 AspSerArgLeuHisSerAlaProMetTyrPhePheLeuSerHisLeuAla 67
204 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCGGT 253

 67 lValAspIleAlaTyrAlaCysAsnThrrValProArgMetLeuValAsnL 84
254 CGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACC 303

 84 euLeuHisIleProAlaLysProIleSerSerPheAlaLaGlyArgMetMetGlnThr 100
304 TCCTGCACATCCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 353

101 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuVaI 117
354 TCGTACGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA 403

117 PheLeuPheSerThrPheAlaValThrCysLeuLeuLeuLeuValValMe 134
404 TTCCTATTTTCCACTTTTGCTGTGACCTGCCTCCTCCTGGTGGTGAT 453

134 leMetThrThrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
454 TCATGACCACTCGGAGAGTCTGCATCACCCTGGCGGTGACTTCCTGGACACCACT 503

151 GlyValLeuLeuSerLeuValLeuLeuLeuLeuProLeuProPh 167
504 GGAGTCCTTTTATCCTTGAATTCAATCTTGTGTTACTTCTACCTTTACCTT 553

167 eCysArgProGlnLysIleThrYrHisPhePheCysGluIleLeuAlaVall 184
554 CTGTAGGCCCCAGAAAAATTATCACTTTTTTTGTGAAATCTTGGCTGTTC 603

184 eulyLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
604 TCAAACTTGCCTGTGCAGATACCCACATCAAATGAGAACATGGTCTTTGGCC 653

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
654 GGAGCAAATTGCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATA 703

217 rMetCysIleLeuCysAlaIleIleLeuGlnIleGlnInSerArgGlnValGlnA 234
704 TATGTGCATCCTCTGTGCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGA 753

234 rglyLysAlaPheCysThrCysPheSerHisLeuCyssValLeuGlyLeuPhe 250
754 GGAAAGCCTTCTGCACCTGCTTCTCCCACCTGTGTGATGGAACCCA 803

251 TyrGlyThrAlalleIleMettYrvalGlyProArgTyrGlyAsnProly 267
804 TATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCAA 853
```

seq_name: gb_pr:AC004889

seq_documentation_block:
```
LOCUS           AC004889           130030 bp    DNA      linear   PRI 14-JAN-1999
DEFINITION  Homo sapiens clone DJ0798C17, complete sequence.
ACCESSION   AC004889
VERSION     AC004889.1  GI:4156187
KEYWORDS    HTG.
SOURCE      human.
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 130030)
AUTHORS   Waterston,R.H.
TITLE     The sequence of Homo sapiens clone
JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 130030)
AUTHORS   Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 130030)
AUTHORS   Waterston,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (14-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT   On Jan 14, 1999 this sequence version replaced gi:3212108.
FEATURES             Location/Qualifiers
source          1..130030
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="DJ0798C17"
ORIGIN
BASE COUNT    38117 a   25410 c   26567 g   39936 t
```

alignment_block:
US-09-898-586-24 x AC004889/rev

Align seg 1/1   to reverse of: AC004889   from: 1   to: 130030

alignment_scores:
```
Quality: 1609.00            Length:    310
Ratio: 5.190                Gaps:        0
Percent Similarity: 100.000   Percent Identity: 100.000
```

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
  ||||||||||||||||||||||||||||||||||||||||||||||||||
1 ATGGGGGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTYC 111194

17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
   ||||||||||||||||||||||||||||||||||||||||||||||||||
111193 CGTTGGCCCAAGGATTCAGAGATGCTCCTCTTGTTTGGGCTTCTTCCTGTCT 111144

34 yrValPheThrLeuLeuGlyLeuAsnGlyThrIleLeuGlyLeuIleSerLeu 50
   ||||||||||||||||||||||||||||||||||||||||||||||||||
111143 ACGTCTTCACCCTGCTGGGGACNGGGACCATACTGGGGCTCATCTCACTG 111094

51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
   ||||||||||||||||||||||||||||||||||||||||||||||||||
```

seq_name: gb_pat:AX242191

seq_documentation_block:
```
LOCUS           AX242191           930 bp    DNA      linear   PAT 26-SEP-2001
DEFINITION  Sequence 939 from Patent WO0127158.
ACCESSION   AX242191
VERSION     AX242191.1  GI:15799066
KEYWORDS    
SOURCE      synthetic construct.
ORGANISM  synthetic construct
```

```
301 LysArgValLeuGlyValGluArgAlaLeu 310
    ||||||||||||||||||||||||||||||
110343 AAGAGAGTGCTGGGAGTGAGAAAGGGCTTTA 110314
```

Left column

REFERENCE   artificial sequence.
             1 (bases 1 to 930)
AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
             Yanai,I.
TITLE   Olfactory receptor sequences
JOURNAL   Patent: WO 0127158-A 939 19-APR-2001;
             Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES   Location/Qualifiers
    source   1..930
             /organism="synthetic construct"
             /db_xref="taxon:32630"
             /note="(H38g789 nucleotide)"
BASE COUNT   192 a   266 c   201 g   271 t
ORIGIN

alignment_scores:
    Quality: 1602.00   Length: 310
    Ratio: 5.184   Gaps: 0
    Percent Similarity: 99.677   Percent Identity: 99.677

alignment_block:
US-09-898-586-24 x AX242191

Align seq 1/1 to: AX242191   from: 1 to: 930

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGGGGGACAATATAACATCACACAGAGTTCCTCCTACTGGGATTTCC   50

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT  34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 CGTTGGGCCCAAGGATTCAGATGCTTCCTTCTTTGGGCTTCTTCCCTGTTCT  100

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyIleLeuIleSerLeu  50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 ACGTCTTCACCCTGCTGGGAACGGGGACCATACTGGGCTCATCTACTG  150

 51 AspSerArgLeuHisIleAlaProMetTyrPhePheLeuSerHisLeuAlaV  67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GACTCCAGACTGCACGCCCCCCATGTACTTCTTCCTCACACCTGGCGGT  200

 67 lValAspIleAlaTyrAlaCysAsnThrValIleProArgMetLeuValAsnL  84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 CGTCGACATCGCCTACGCCTGCAACACGGTGCCCCCGATGCTGGTGAACC  250

 84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr  100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 TCCTGCACCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC  300

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValMe  117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 TTTCTGTTTTCCACTTTTGCTGTCACAGAAATGTCTCCTCCTGTGGTGAT  350

117 tSerTyrAspLeuTyrValAlaIleCysTyrHisProLeuArgTyrLeuAlaI  134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 GTCCTATGACCTGTACGTGGCCATCTGCCACCCCCTCCGAATATTTGGGCA  400

134 leMetThrTrpArgValCysSerIleThrLeuAlaValThrSerTrpThrThr  150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 TCATGACACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGGACCACT  450

151 GlyValLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh  167
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 GGAGTCCTTTTATCCTTGATTCATCTTGTTGTTACTTCCACTTTACCCTT  500

167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuLeuAlaValI  184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 CTGTAGGCCCCAGAAAATTTACACTTTTTTGTGAAATCTTGGCTGTTC  550

184 euLysLeuAlaCysSerAlaAspThrIleAsnGluAsnMetValLeuAla  200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 TCAAACTTGGCTGTGCAGATACCCACATCAATGAGAACATGGTCTGGGCC  600
```

Right column

```
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTyr  217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 GGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATA  650

217 rMetCysIleLeuCysSerAlaIleLeuGlnIleLeuSerArgGluValGlnA  234
    ||||||||||||||||||||||||||||||||||||||||||||||||||
651 TATGTGCATCCCTCTGCCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA  700

234 rglysAlaPheCysThrCysPheSerHisLeuCysSerValIleGlyIleuPhe  250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 GGAAAGCCTTCTCGCACCTGCTTCTCCCACCTCTGTCTGTAATAGGAATCT  750

250 TyrGlyThrAlaIleLeuMetTyrrValGlyProArgTyrGlyAsnProLy  267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 TATGGCCAGCCCAATTATACATGTAATGTTGGACCCAGATTATGGGAACCCAA  800

267 sGluGlnIleLysSerTyrLeuLeuPheHisSerSerLeuPheAsnProMetL  284
    ||||||||||||||||||||||||||||||||||||||||||||||||||
801 GGAGCAGAAGAAATATCTCCCGCTGTTTCACAGCCTCTTAATCCCATGC  850

284 euAsnProLeuIleCysSerLeuArgAsnSerGluValGlylysAsnThrLeu  300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 TCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTTGAAGAANACTTTG  900

300 LysArgValLeuGluGlyArgAlaLeu  310
    |||||||||||||||||||||||
901 AAGAGAGTGCTGGGAGTTAGAAAAGGGCTTTTA  930
```

seq_name: gb_pat:AX241442

seq_documentation_block:
LOCUS   AX241442   930 bp   DNA   linear   PAT 26-SEP-2001
DEFINITION Sequence 190 from Patent WO0127158.
ACCESSION   AX241442
VERSION   AX241442.1 GI:15798317
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 930)
  AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
            Yanai,I.
  TITLE   Olfactory receptor sequences
  JOURNAL   Patent: WO 0127158-A 190 19-APR-2001;
            Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES   Location/Qualifiers
    source 1..930
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="(H38g38 nucleotide)"
BASE COUNT   193 a   265 c   202 g   270 t
ORIGIN

alignment_scores:
    Quality: 1584.00   Length: 310
    Ratio: 5.160   Gaps: 0
    Percent Similarity: 99.032   Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AX241442

Align seq 1/1 to: AX241442   from: 1 to: 930

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGGGGGACAATATAACATCACACAGAGTTCCTCCTACTGGGATTTCC   50

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT  34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 CGTTGGGCCCAAGGATTCAGATGCTTCCTTCTTTGGGCTTCTTCCCTGTTCT  100
```

seq_name: gb_pat:AX241473

seq_documentation_block:
LOCUS   AX241473   930 bp   DNA   linear   PAT 26-SEP-2001
DEFINITION   Sequence 221 from Patent WO0127158.

ACCESSION   AX241473
VERSION   AX241473.1   GI:15798348
KEYWORDS   .
SOURCE   synthetic construct.
ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 930)
AUTHORS   Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
             Yanai,I.
TITLE   Olfactory receptor sequences
JOURNAL   Patent: WO 0127158-A 221 19-APR-2001;
             Digiscents (US); YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES   Location/Qualifiers
     source   1..930
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="(H38970 nucleotide)"
BASE COUNT   193 a   265 c   202 g   270 t
ORIGIN

alignment_scores:
          Quality: 1584.00   Length: 310
          Ratio: 5.160   Gaps: 0
          Percent Similarity: 99.032   Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AX241473

Align seq 1/1 to: AX241473   from: 1 to: 930

seq_name: gb_pat:AX197438

seq_documentation_block:

LOCUS	AX197438	1040 bp	DNA	linear	PAT 29-AUG-2001

DEFINITION Sequence 3 from Patent WO0151632.
ACCESSION   AX197438
VERSION   AX197438.1 GI:15387828
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1040)
  AUTHORS   Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tcherney,V.T.,
            Spytek,K.A. and Li,L.
  TITLE   Odorant receptor polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0151632-A 3 19-JUL-2001;
            Curagen Corporation (US)
FEATURES   Location/Qualifiers
     source   1..1040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   229 a   285 c   221 g   305 t
ORIGIN

```
  501 CTGTAGGCCCCAGAAAAATTTATCACTTTTTTGTGAAATCTTGGCTGTTC 550
        |||||||||||||||||||||||||||||||||||||||||||||||||
  184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
  551 TCAAACTTGCCTGTGCAGANTACCCACATCANTGAGAACATGGTCTTGGCC 600
        |||||||||||||||||||||||||||||||||||||||||||||||||
  201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
  601 GGAGCAANTTTCTGGGCTGGGTGGGACCCTTGTCCACAATTGTAGTTTCATA 650
        |||||||||||||||||||||||||||||||||||||||||||||||||
  217 rMetCysIleLeuCysValAlaIleLeuGlnLeuSerArgGluValGlnA 234
  651 TATGTCCATCCTCTGTGTTNCACAGCTCTGAATCCAATCAAAGGGAAGTTCAGA 700
        |||||||||||||||||||||||||||||||||||||||||||||||||
  234 rgLysAlaPheCysThrCysPheSerHisLeuCysValValIleGlyLeuPhe 250
  701 GGAAAAGCCTTCCGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCGTT 750
        |||||||||||||||||||||||||||||||||||||||||||||||||
  251 TyrGlyThrAlaIleIleMetGlyArgTyrGlyAsnProLy 267
  751 TATGGCACAGCCATTATCATGTATGTTGGACCCAGATPATGGGAACCCAA 800
        |||||||||||||||||||||||||||||||||||||||||||||||||
  267 sGluGlnLnLysLysTyrIleLeuLeuPheHisSerLeuLeuPheHisAsnProMetL 284
  801 GGAGCAGAAGAAAATATCTCCTGCTGTTTCACAGCCTCTTNAATCCCATGC 850
        |||||||||||||||||||||||||||||||||||||||||||||||||
  284 euAsnProLeuIleCysSerLeuArgAsnSerGluValValLysAsnThrLeu 300
  851 TCAATCCCCTTATCTGTACTCTTNAGGAACTCAGAAGTGAAGAATACTTTG 900
        |||||||||||||||||||||||||||||||||||||||||||||||||
  301 LysArgValLeuGlyValGluArgAlaLeu 310
  901 AAGAGAGTGCTCGGGAGTAGAAAAGGGCTTTA 930
        ||||||||||||||||||||||||||||||||
```

alignment_scores:
          Quality: 1584.00
          Ratio: 5.160
          Percent Similarity: 99.032   Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AX197438

Align seg 1/1   to: AX197438   from: 1   to: 1040

```
    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   82 ATGGGAGACAATATAACATCCATCANGAGAGTTCCTCCTACTCGGATTTCC 131
```

```
   17 oValGlyProArgIleGlnMetLeuLeuPheSerLeuPheI 34
        |||||||||||||||||||||||||||||||||||||||||
  132 CGTTGGCCCAAGGATTCAGATGCTTCCTCTTTCCCCGTTCT 181
   34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  182 ACGTCTTCACCCTGCTGGGAAACGGGACCATACTGGGGCTCATCTCACTG 231
   51 AspSerArgLeuHisSerAlaProMetTyrPhePheLeuGlyHisLeuAla 67
        |||||||||||||||||||||||||||||||||||||||||||||||||
  232 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCCACCTGGGCGGGT 281
   67 lValAspIleAlaIleTyrAlaCysAsnThrValProArgMetLeuAsnL 84
        |||||||||||||||||||||||||||||||||||||||||||||||||
  282 CGTCCAGATCGCCATCTACGCCTGCAACACGGTGCCCCGGATGCTGAACC 331
   84 euLeuHisProAlaLysSerProIleSerArgPheAlaGlyArgMetMetGlnThr 100
        |||||||||||||||||||||||||||||||||||||||||||||||||
  332 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGGCCGCARGAGATGCAGACC 381
  101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  382 TTTCTGTTCTCCACTTTTGCCGTCACAGAAAGTCTCCTCCTGGTGGTGAT 431
  117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
        |||||||||||||||||||||||||||||||||||||||||||||||||
  432 GTCCTATGANCTCTACGTGGCCATCTGCCACCCCCTCCGATAATTTGGCCA 481
  134 leMetThrTrpArgValCysIleThrLeuAlaAlaLThrSerTrpThrTr 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  482 TCATGACCTGGAGAGTCTGCATCACCCTCGCCGGTGACTTCCTGGACCACT 531
  151 GlyValLeuLeuSerLeuIleHisIleSerValLeuLeuLeuProLeuProPh 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  532 GGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTTTACCCTT 581
  167 eCysArgProGlnIleTyrHisPhePheCysGluIleLeuAlaLeuAvalI 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  582 CTGTAGGCCCCCAGAAAAAATTTATCACTTTTTTTTTGTGAAAATCTTGGTTTC 631
  184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  632 TCAAACTTGCCTGTGCAGANTACCCACATCAATGAGAACATGGTCTTGGCC 681
  201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  682 GGAGCAAATTCTGGGCTGGTGGGACCCTTGTCCACAATTGTCATA 731
  217 rMetCysIleLeuCysValAlaIleLeuGlnLeuSerArgGluValGlnA 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  732 TATGTGCATCCTCTGTGTTACACAGCTCTGAATCCAATCAAAGGGAAGTTCAGA 781
  234 rgLysAlaPheCysThrCysPheSerHisLeuCysValValIleGlyLeuPhe 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
  782 GGAAAGCCTTCCGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCGTT 831
  251 TyrGlyThrAlaIleIleMetGlyArgTyrGlyAsnProLy 267
        |||||||||||||||||||||||||||||||||||||||||
  832 TATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCAA 881
  267 sGluGlnLnLysLysTyrIleLeuLeuPheHisSerLeuLeuPheHisAsnProMetL 284
        |||||||||||||||||||||||||||||||||||||||||||||||||
  882 GGAGCAGAAGAAAATATCTCCTGCTGTTTCACAGCCTCTTNAATCCCATGC 931
  284 euAsnProLeuIleCysSerLeuArgAsnSerGluValValLysAsnThrLeu 300
        |||||||||||||||||||||||||||||||||||||||||||||||||
  932 TCAATCCCCTTATCTGTACTCTTAGGAACTCAGAAGTGAAGAATACTTTG 981
  301 LysArgValLeuGlyValGluArgAlaLeu 310
        ||||||||||||||||||||||||||||||
  982 AAGAGAGTGCTCGGGAGTAGAAAAGGGCTTTA 1011
```

seq_name: gb_pat:AX197469

seq_documentation_block:
LOCUS   AX197469   1040 bp   DNA   linear   PAT 29-AUG-2001
DEFINITION   Sequence 34 from Patent WO0151632.
ACCESSION   AX197469
VERSION   AX197469.1 GI:15387841
KEYWORDS   .
SOURCE   human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1040)
  AUTHORS   Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,
             Spytek,K.A. and Li,L.
  TITLE   Odorant receptor polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0151632-A 34 19-JUL-2001;
             Curagen Corporation (US)
FEATURES   Location/Qualifiers
     source   1..1040
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   229 a   285 c   221 g   305 t
ORIGIN

alignment_scores:
     Quality: 1584.00
     Ratio: 5.160
     Percent Similarity: 99.032   Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AX197469

Align seq 1/1 to: AX197469 from: 1 to: 1040

                  1 to: AX197469 from: 1 to: 1040

    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
      |||||||||||||||||||||||||||||||||||||||||||||||||
   82 ATGGGAGACAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTCC 131

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
      |||||||||||||||||||||||||||||||||||||||||||||||||
  132 CGTTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCT 181

   34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
      |||||||||||||||||||||||||||||||||||||||||||||||||
  182 ACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCAGCACTG 231

   51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
      |||||||||||||||||||||||||||||||||||||||||||||||||
  232 GACTCCAGACTGCACGCCCCATGTACTTCTTCCTGTCTCACCTGGCGGT 281

   67 lValAspSerLeuArgTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
      |||||||||||||||||||||||||||||||||||||||||||||||||
  282 CGTCGACAGTCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACC 331

   84 euLeuHisProAlaLeuSerProPheAlaLeuGlyArgMetMetGlnThr 100
      |||||||||||||||||||||||||||||||||||||||||||||||||
  332 TCCTGCATCCAGCCACAGCCATCTCCTTTGCGCGGCCCCATGCAGACC 381

  101 PheLeuPheSerThrPheAlaValThrGluCysGlyLeuLeuValAlaMe 117
      |||||||||||||||||||||||||||||||||||||||||||||||||
  382 TTTCTGTTTTCCACTTTTCCTGTCACAGAATGTCTCCCTCCTGGTGATT 431

  117 tSerTyrArgAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAI 134
      |||||||||||||||||||||||||||||||||||||||||||||||||
  432 GTCCTATGATCTGTACGTGGCCAACTGCCACCCCCTCCGATATTTGGCCA 481

  134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
      ||||||||||||||||||||||||||||||||||||||||||||||
  482 TCATCACCTGGAGAGTCTGCATCACCCTGGCGGTGACTTCCTGGACCACT 531

  151 GlyValLeuLeuSerLeuHisLeuValLeuLeuLeuProLeuProh 167
      ||||||||||||||||||||||||||||||||||||||||||||||
  532 GGAGTCCTTTTATCCCTTGATTCAGTTGTTGTTACTTCCTTTACCCTT 581

  167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVaIlI 184
      |||||||||||||||||||||||||||||||||||||||||||||||||
  582 CTTGTAGGCGCCCCAGAAAATTATCACTTTTTTGTGAAATCTTGGCTGTTC 631

  184 euLysLeuAlaCysAlaAspThrHisIleLeuAsnGluAsnMetValLeuAla 200
      ||||||||||||||||||||||||||||||||||||||||||||||||||
  632 TCAAACTTGCCTGCTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 681

  201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrLeuValValSerTy 217
      |||||||||||||||||||||||||||||||||||||||||||||||||
  682 GGAGCAATTTCTGGGCGTGGTGGGACCCTTGTCCACAATTGTAGTTTCATA 731

  217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
      |||||||||||||||||||||||||||||||||||||||||||||||||
  732 TATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 781

  234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
      |||||||||||||||||||||||||||||||||||||||||||||||||
  782 GGAAAGCCTTCCCGCACCTGCTTCTCCCACCTCTGTGTCGAATGGACTCGTT 831

  251 TyrGlyThrAlaIleIleMetThrTyrValGlyProArgTyrGlyAsnProLy 267
      |||||||||||||||||||||||||||||||||||||||||||||||||
  832 TATGGCACAGCCCATTATCATGTATGTTGGACCCAGAATATGGGAACCCCAA 881

  267 sGluIleGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
      |||||||||||||||||||||||||||||||||||||||||||||||||
  882 GGAGCAGAAGAAAATATCTCCTTGTTTCACGGCTCTTTAATCCCATGC 931

  284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
      |||||||||||||||||||||||||||||||||||||||||||||||||
  932 TCAATCCCCTTATCTGTAGTCTTAGACAACTCAGAACGTGAAGAATACTTTG 981

  301 LysArgValLeuGlyValGlyLuArgAlaLeu 310
      |||||||||||||||||||||||||||
  982 AAGAGAGTGCTGGGAGTAGAAAAGGGCTTTA 1011

seq_name: gb_pri:AL135904

seq_documentation_block:
LOCUS   AL135904   59130 bp   DNA   linear   PRI 20-JUL-2000
DEFINITION   Human DNA sequence from clone RP5-1005H11 on chromosome 6 Contains
             part of the gene for a 7 transmembrane receptor (rhodopsin family)
             (olfactory receptor like) protein, the gene for WUSGC:H.DJ0988G15.3
             protein, part of the PDNP3 (phosphodiesterase I/nucleotide
             pyrophosphatase 3) gene, ESTs, STSs and GSSs, complete sequence.
ACCESSION   AL135904
VERSION   AL135904.11 GI:7159399
KEYWORDS   HTG.
SOURCE   human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 59130)
  AUTHORS   Lovell,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (05-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
             requests: clonerequest@sanger.ac.uk
COMMENT   On Mar 6, 2000 this sequence version replaced gi:7105768.
             During sequence assembly data is compared from overlapping clones.
             Where differences are found these are annotated as variations
             together with a note of the overlapping clone name. Note that the
             variation annotation may not be found in the sequence submission
             corresponding to the overlapping clone, as we submit sequences with
             only a small overlap as described above.
             This sequence has been finished according to sequence map criteria
             as follows. An attempt is made to resolve all sequencing problems,
             such as compressions and repeats, but not necessarily within known
             annotated human repeat sequence elements (e.g. Alu). Where the
             sequence is ambiguous, there is an annotation using the 'unsure'

```
           feature key.
           The following abbreviations are used to associate primary accession
           numbers given in the feature table with their source databases:
           EMi, EMBLi Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
           on the WORMPEP database can be found at
           http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
           was generated from part of bacterial clone contigs of human
           chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
           Group. Further information can be found at
           http://www.sanger.ac.uk/HGP/chr6
           RP5-1005H11 is from the library RPCI-5 constructed at the Roswell
           Park Cancer Institute by the group of Pieter de Jong. For further
           details see http://bacpac.med.buffalo.edu/
           VECTOR: pCYPAC2
           IMPORTANT: This sequence is not the entire insert of clone
           RP5-1005H11 It may be shorter because we sequence overlapping
           sections only once, except for a 100 base overlap.
           The true left end of clone RP5-1005H11 is at 59130 in this
           sequence. The true right end of clone RP5-914N13 is at 105 in this
           sequence.

FEATURES            Location/Qualifiers
source              1..59130
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="6"
                    /clone="RP5-1005H11"
                    /clone_lib="RPCI-5"
repeat_region       1..68
                    /note="Alus repeat: matches 1..68 of consensus"
repeat_region       546..844
                    /note="AluJo repeat: matches 2..301 of consensus"
repeat_region       870..940
                    /note="U4 repeat: matches 1..72 of consensus"
repeat_region       1691..2906
                    /note="L2 repeat: matches 1425..2737 of consensus"
repeat_region       3457..3960
                    /note="HAL1 repeat: matches 420..971 of consensus"
repeat_region       3971..4106
                    /note="MLT1D repeat: matches 1..129 of consensus"
repeat_region       4249..4282
                    /note="17 copies 2 mer tg 100% conserved"
repeat_region       4966..5147
                    /note="AluJb repeat: matches 1..182 of consensus"
repeat_region       5216..5307
                    /note="L2 repeat: matches 2405..2503 of consensus"
repeat_region       5495..5794
                    /note="AluSx repeat: matches 1..296 of consensus"
repeat_region       5891..6180
                    /note="AluSx repeat: matches 1..291 of consensus"
repeat_region       7607..7618
                    /note="L1M1 repeat: matches 4856..4867 of consensus"
repeat_region       7619..8347
                    /note="L1PA16 repeat: matches 5421..6156 of consensus"
repeat_region       8348..8898
                    /note="L1M1 repeat: matches 4866..5403 of consensus"
repeat_region       9018..9866
                    /note="AluSx repeat: matches 1..836 of consensus"
repeat_region       9867..10189
                    /note="TIGGER1 repeat: matches 1..312 of consensus"
repeat_region       10190..11723
                    /note="TIGGER1 repeat: matches 836..2418 of consensus"
repeat_region       12404..12715
                    /note="AluSx repeat: matches 1..312 of consensus"
misc_feature        complement(13947..14176)
                    /note="match: GSS: Em:B74942"
repeat_region       14280..14383
                    /note="AluSg/x repeat: matches 190..294 of consensus"
repeat_region       15340..15508
                    /note="L2 repeat: matches 2563..2750 of consensus"
repeat_region       15584..15968
                    /note="AluSx repeat: matches 35..312 of consensus"
repeat_region       16527..16612
                    /note="MIR repeat: matches 50..138 of consensus"
```

```
repeat_region       16731..17027
                    /note="AluSg repeat: matches 1..295 of consensus"
repeat_region       17609..17747
                    /note="MER91A repeat: matches 1..153 of consensus"
repeat_region       17867..18162
                    /note="AluY repeat: matches 1..299 of consensus"
repeat_region       18461..18950
                    /note="MER1A repeat: matches 2..527 of consensus"
repeat_region       19185..19224
                    /note="20 copies 2 mer tg 85% conserved"
misc_feature        19474..19712
                    /note="match: STS: Em:G16149"
repeat_region       20037..20361
                    /note="AluSx repeat: matches 1..311 of consensus"
gene                complement(21566..22384)
                    /gene="dJ1005H11.1"
CDS                 complement(21566..>22384)
                    /gene="dJ1005H11.1"
                    /note="match: proteins: Sw:P47881 Tr:070265 Tr:070266
                    Tr:070267 Tr:070268 Tr:062942 Tr:095918 Tr:095047"
                    /codon_start=1
                    /evidence=not_experimental
                    /product="dJ1005H11.1 (7 TRANSMEMBRANE RECEPTOR (RHODOPSIN
                    FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN))"
                    /protein_id="CAB99212.1"
                    /db_xref="GI:9369991"
                    /translation="LGNGTILGLISLDSRLHAPMYFFLSHLAVVDIAYACNTVPRMLV
                    NLLHPAKPISFAGRMMQTFLFSTFAVTECLLLVVMSYDLYVAICHPLRYLAIMTWRVC
                    ITLAVTSWFTGVLLSILHLVLLPLPFCPREQKIYHFFCEILAVLKLACADTHINENMV
                    LAGAISGLVGPLSTIVVSYMCIICAILQIOSREVQRKAPRTCFSHLCVIGLYYGTAII
                    MVVGPRYCGNPKEQKYLLHFSLFNPMLNPLICSLRNSEVKNTLKRVLGVERAL"
repeat_region       22385..22489
                    /note="L2 repeat: matches 2461..2551 of consensus"
repeat_region       22696..23011
                    /note="AluSx repeat: matches 1..312 of consensus"
misc_feature        23044..23513
                    /note="match: STS: Em:H39853"
misc_feature        complement(23123..23716)
                    /note="match: GSS: Em:AI065560"
misc_feature        complement(23300..23543)
                    /note="match: GSS: Em:AQ276140"
misc_feature        23751..24305
                    /note="match: GSS: Em:AQ058777"
misc_feature        23753..24176
                    /note="match: GSS: Em:AQ798220"
misc_feature        23762..24465
                    /note="match: GSS: Em:AQ550610"
misc_feature        24220..24343
                    /note="match: GSS: Em:B86674"
repeat_region       24220..24343
                    /note="FLAM_C repeat: matches 1..124 of consensus"
misc_feature        complement(24681..253395)
                    /note="match: GSS: Em:AQ627093"
misc_feature        complement(24780..25383)
                    /note="match: GSS: Em:AQ627184"
repeat_region       26053..26548
                    /note="AluSx repeat: matches 1..521 of consensus"
repeat_region       26549..26861
                    /note="TIGGER1 repeat: matches 1..302 of consensus"
repeat_region       26865..27121
                    /note="L1M4 repeat: matches 4025..4339 of consensus"
repeat_region       27122..27424
                    /note="Tigger3b repeat: matches 2..272 of consensus"
repeat_region       27425..27579
                    /note="AluSx repeat: matches 272..451 of consensus"
repeat_region       27582..27658
                    /note="Tigger3b repeat: matches 1155..1231 of consensus"
repeat_region       27659..29067
                    /note="Tigger3b repeat: matches 1..1155 of consensus"
repeat_region       29189..29287
                    /note="L1M4 repeat: matches 2585..2600 of consensus"
repeat_region       29916..29999
                    /note="L1M4 repeat: matches 2499..2600 of consensus"
misc_feature        29916..29999
                    /note="match: STS: Em:AA931824"
```

```
misc_feature    29507..29990
                /note="match: STS: Em:AA931824"
misc_feature    29530..29981
                /note="match: STS: Em:AQ356766"
misc_feature    29530..29737
                /note="match: GSS: Em:AQ253729"
misc_feature    29533..29855
                /note="match: GSS: Em:AQ319835"
misc_feature    29938..29943
                /note="match: GSS: Em:AA235111"
misc_feature    29560..29787
                /note="match: STS: Em:G27651 Em:T99793"
gene            complement(29779..32112)
                /gene="dJ1005H11.2"
CDS             complement(29779..32112)
                /gene="dJ1005H11.2"
                /note="match: proteins: Tr:O95046"
                /codon_start=1
                /evidence=not_experimental
                /product="dJ1005H11.2 (WUGSC:H_DJ0988G15.3 PROTEIN)"
                /protein_id="CAB99213.1"
                /db_xref="GI:9368992"
                /translation="MEEPGATPQPYLGLVLEELGRVVAALPESMRPDENPYGFPSELV

alignment_scores:
              Quality:     1584.00          Length:       310
              Ratio:         5.160          Gaps:           0
Percent Similarity:       99.032   Percent Identity:     99.032

alignment_block:
US-09-898-586-24 x Al135904/rev

Align seg 1/1    to reverse of: Al135904     from: 1    to: 59130
```

```
    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheThr    17
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22498 ATGGGAGACAATATAACATCCATACAGAGAGTTCTCCTACTGGGATTTCC  22449

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT    34
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22448 CGTTGGCCCAAGGATTCAGAGAGCCTCCTCTTCTTGGCTCTTCTCCCTGTTCT  22399

   34 yrValPheThrHisLeuLeuGlyAsnGlyThrHisLeuGlyLeuIleSerLeu  50
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22398 ACGTCTTCACCCCTGCTGGGGAACGGGACACATACTGGGGCTCATCTCACTG  22349

   51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa    67
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22348 GACTCCAGACTGCACGCCCCCATGTACTTCTTCTTTCCTCACACTGGCGT  22299

   67 IValAspIleAlaLTyrAlaCysAsnThrValProArgMetLeuValAsnL    84
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22298 CGTCGACAATCGCCTACGCCTGCAACACGGTGCCCGGAGATGCTGGTGAACC  22249

   84 euLeuHisProAlaLysProLeuSerPheAlaGlyArgMetMetGlnThr   100
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22248 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCGCCATGATGCAGACC  22199

  101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValIleMe   117
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22198 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGTGAT  22149

  117 tSerTyrTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI   134
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22148 GTCCTATTACGATCTGTACGTGGCCATCTGCCACCCCCTCAGATATTTGGGCCA  22099

  134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr   150
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22098 TCATGACCTGGAGAGTCTGCATCACCCTCGCGGTCACTTCCTGGACCACT  22049

  151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh   167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
22048 GGAGTCCTTTTTATCCTTGTGTTACTTCTACCTTTACCCTT  21999
```

```
  167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuLeuAlaVal   184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21998 CTGTAGGCCCCAGAAAATTTATCACTTTTTTTGTGAAATCTTTGGCTGTTC  21949

  184 euLLysLeuAlaCysAlaAspThrHisIleLeuAsnGluAsnMetValLeuA   200
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21948 TCAAACTTGCCTGTCGTGCAGATACCACATCAATGAGAACATGGTCTTGGCC  21899

  201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy   217
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21898 GGAGCAAATTTCTGGGCTGGTGGGACCCTTGTCCACAAATTGTGAGTTCATA  21849

  217 rMetCysIleLeuCysAlaIleIleLeuGlnSerArgGluValGlnA   234
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21848 TAHGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGAGTTCAGA  21799

  234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe   250
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21798 GGAAAGCCTTCCGCACCTGCTTCTCCCACCTCTGTGTCATAGGACTCGTT  21749

  251 TyrGlyThrAlaIleIleLeMetThrTyrValGlyProArgTyrGlyAsnProLy   267
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21748 TAHGGCACAGCCATTATCATGTCATATGTTGGACCCCAGATATGGGAACCCAA  21699

  267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetL   284
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21698 GGAGCAGAAGAAAATATCTCCTGCTGTTTCACAGCCTCTTAATCCCATGG  21649

  284 euAsnProLeuIleCysSerSerLeuArgAsnSerGlyValLysAsnThrLeu   300
      ||||||||||||||||||||||||||||||||||||||||||||||||||
21648 TCAAATCCCCTTATCTGTATGTCTTAGGAAACTTCAGAAGTGAAGAAATACTTTG  21599

  301 LysArgValLeuGlyValGlyArgAlaLeu   310
      ||||||||||||||||||||||||||||||
21598 AAGAGAGTGCTCGGGAGTAGAAGGGCTTTA  21569
```

```
seq_name: gb_pr:AC005587

seq_documentation_block:
LOCUS       AC005587      140915 bp    DNA    linear    PRI 21-DEC-1999
DEFINITION  Homo sapiens PAC clone RP5-988G15 from 7q33-q35, complete sequence.
ACCESSION   AC005587
VERSION     AC005587.1    GI:4156166
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 140915)
AUTHORS     Hou,S., Wohldmann,P. and Le,T.
TITLE       The sequence of Homo sapiens PAC clone RP5-988G15
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 140915)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (01-SEP-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 140915)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (14-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 140915)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (15-JAN-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 140915)
AUTHORS     Waterston,R.
TITLE       Direct Submission
```

```
JOURNAL     Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Jan 14, 1999 this sequence version replaced gi:3907499.
COMMENT     ------------------
            Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc
              Contact: sapiens@watson.wustl.edu
            ------------------
            Summary Statistics
              Center project name: H_DJ0988G15
            ------------------

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry; an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one subclone; and the assembly was
            confirmed by restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7 ; send
            mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-5, prepared by
            Pieter de Jong and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu) using the method described by
            Ioannou et al., Nature Genetics 6:84-9 (1994).  The library is from
            one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
            VECTOR: pCYPAC2
            NEIGHBORING SEQUENCE INFORMATION:
            Actual start of this clone is at base position 1 of RP5-988G15;
            actual end is at 140915 of RP5-988G15.
FEATURES             Location/Qualifiers
     source          1..140915
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q33-q35"
                     /clone="RP5-988G15"
                     /clone_lib="RPCI-5"
     repeat_region   13..2090
                     /rpt_family="L1"
     repeat_region   2097..2299
                     /rpt_family="L1"
     repeat_region   2300..2603
                     /rpt_family="L1"
     repeat_region   2604..4120
                     /rpt_family="Alu"
     repeat_region   4123..4166
                     /rpt_family="L1"
     repeat_region   4168..4307
                     /rpt_family="(TA)n"
     repeat_region   6635..6668
                     /rpt_family="L1"
     repeat_region   6829..7260
                     /rpt_family="AT_rich"
     repeat_region   7261..7642
                     /rpt_family="L1"
     repeat_region   /rpt_family="MaLR"

     repeat_region   7643..7923
                     /rpt_family="L1"
     repeat_region   7982..8283
                     /rpt_family="L1"
     repeat_region   8286..8412
                     /rpt_family="Alu"
     repeat_region   8413..8708
                     /rpt_family="L1"
     repeat_region   8709..8838
                     /rpt_family="Alu"
     repeat_region   8840..8886
                     /rpt_family="L1"
     repeat_region   9111..9139
                     /rpt_family="MaLR"
     repeat_region   9204..9490
                     /rpt_family="AT_rich"
     repeat_region   9646..9924
                     /rpt_family="Alu"
     repeat_region   10286..10455
                     /rpt_family="Alu"
     repeat_region   10759..10838
                     /rpt_family="L2"
     repeat_region   11127..11266
                     /rpt_family="Mariner"
     repeat_region   11850..12154
                     /rpt_family="L2"
     repeat_region   12172..12211
                     /rpt_family="Alu"
     repeat_region   12352..12469
                     /rpt_family="AT_rich"
     repeat_region   12458..12607
                     /rpt_family="(TACG)n"
     repeat_region   12580..12733
                     /rpt_family="(CATA)n"
     repeat_region   12660..12777
                     /rpt_family="(TA)n"
     repeat_region   12778..13031
                     /rpt_family="(CATA)n"
     repeat_region   13226..13535
                     /rpt_family="Alu"
     repeat_region   15279..15520
                     /rpt_family="L2"
     repeat_region   15954..16022
                     /rpt_family="MIR"
     repeat_region   16023..16332
                     /rpt_family="MaLR"
     repeat_region   16333..16371
                     /rpt_family="MER2_type"
     repeat_region   16630..16742
                     /rpt_family="MaLR"
     repeat_region   16816..17068
                     /rpt_family="MaLR"
     repeat_region   17088..17348
                     /rpt_family="L1"
     repeat_region   17857..17979
                     /rpt_family="L1"
     repeat_region   18014..18506
                     /rpt_family="Alu"
     repeat_region   18707..19081
                     /rpt_family="MaLR"
     repeat_region   19135..20717
                     /rpt_family="MaLR"
     repeat_region   20728..20748
                     /rpt_family="Retroviral"
     repeat_region   21444..21599
                     /rpt_family="AT_rich"
     repeat_region   21688..22279
                     /rpt_family="MIR"
     misc_feature    22506..22540
                     /note="60S ribosomal protein L15 pseudogene"
     repeat_region   22557..22583
                     /rpt_family="AT_rich"
     repeat_region   23139..23162
                     /rpt_family="AT_rich"
```

repeat_region   23696..24012   /rpt_family="(TAA)n"
repeat_region   24088..24128   /rpt_family="Alu"
repeat_region   24827..24849   /rpt_family="AT_rich"
repeat_region   24850..25147   /rpt_family="AT_rich"
repeat_region   25148..25173   /rpt_family="Alu"
repeat_region   25951..25974   /rpt_family="AT_rich"
repeat_region   26066..26144   /rpt_family="(CAAAA)n"
repeat_region   26388..26653   /rpt_family="Alu"
repeat_region   27148..27600   /rpt_family="L1"
repeat_region   28002..28294   /rpt_family="L2"
repeat_region   28002..28294   /rpt_family="Alu"

alignment_scores:
   Quality: 1584.00   Length: 310
   Ratio: 5.160   Gaps: 0
   Percent Similarity: 99.032   Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AC005587

Align seg 1/1 to: AC005587 from: 1 to: 140915

1   MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr   17
101599 ATGGGAGACAATATAACAATCATCAGAGAGTTCCTCCTACTGGGATTCC 101648

17   oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT   34
101649 CGTTGGCCCAAGGAATTCAGATAGCTCCTCTTTGGGCTCTTCTCCGTTCT 101698

34   yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu   50
101699 ACGTCTTCACCCTGCTGGGGAACGGGACCATAACTGGGGTCATCTCACTG 101748

51   AspSerArgLeuHisAlaAlaProCysMetTyrPhePheLeuSerHisLeuAla   67
101749 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCGGT 101798

67   lValAspIleAlaTyrAlaCysSerAsnThrValProArgMetLeuValAsnL   84
101799 CGTTGCAGATCGCCTACGCCTGCAAACAGGTGCCCCGGATGCTGGTGAACC 101848

84   euLeuHisIleSerProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr   100
101849 TCCTGCACCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 101898

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe   117
101899 TTTCTGTTTTCCACTTTTGCTGTCACAGAGTGTCTCCTCCTGGTGGTGAT 101948

117 tSerTyrTyrAspLeuGlyTyrValAlaIleCysHisProLeuArgTyrLeuAlaI   134
101949 GTCCTATGATCGTGACTGGGCCATCTGCCACCCCCTGCGATAATTTGGGCCA 101998

134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpPhrThr   150
101999 TCATGACCTGGAGAGTGTCATCACCCTCGCGGTGACTTCCTGGACCACT 102048

151 GlyValLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProH   167
102049 GGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCTACCTTCCCTT 102098

167 eCysArgArgProGlnLysIleTyrHisPhePheCysGluIleLeuLeuAlaValI   184

seq_name: gb_pat:AX197458

seq_documentation_block:
LOCUS   AX197458   1014 bp   DNA   linear   PAT 29-AUG-2001
DEFINITION   Sequence 23 from Patent WO0151632.
ACCESSION   AX197458
VERSION   AX197458.1   GI:15387838
KEYWORDS   .
SOURCE   human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1014)
  AUTHORS   Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,
             Spytek,K.A. and Li,L.
  TITLE   Odorant receptor polypeptides and nucleic acids encoding same
  JOURNAL   Patent: WO 0151632-A 23 19-JUL-2001;
             Curagen Corporation (US)
FEATURES   Location/Qualifiers
     source   1..1014
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   218 a   281 c   216 g   299 t
ORIGIN

alignment_scores:
   Quality: 1568.00   Length: 311
   Ratio: 5.074   Gaps: 1
   Percent Similarity: 99.357   Percent Identity: 99.357

alignment_block:
US-09-898-586-24 x AX197458

Align seg 1/1 to: AX197458 from: 1 to: 1014

1   MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr   17

102099 CTGTAGGCCCCAGAAAAATTATCACTTTTTTTTGTGAAATCTTGGCTGTTC 102148
184   euLysLeuAlaCysSerAlaAspThrHisIleAsnGluAsnMetValLeuuAla   200
102149 TCAAACTTGCCCTGTGCAGATACCCCACATCAAATGAGAACAATGGTCTTTGGCC 102198
201   GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy   217
102199 GGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTTCATA 102248
217   rMetCysIleLeuLeuCysAlaLalIleLeuGlnIleSerArgGluValGlnA   234
102249 TATGTGCATCCTCTGTGCTAATCCTTCAGATCCAATCAACAAGGGAAGTTCAGA 102298
234   rgLysAlaPheCysThrCysPheProSerHisLeuCysValIleGlyLeuPhe   250
102299 GGAAAGCCTTCCCCACCTGCTTCTCCCACCTCGTGTGTGATTGGGACTCGTT 102348
251   TyrGlyThrAlaIleIleIleLeuMetTyrrValGlyProArgTyrGlyAsnProLy   267
102349 TATGGCACAGCCAATTACCAATGTATATGTTGGACCCAGATATGGGAAACCCAA 102398
301   LysArgValLeuGluGlyValGlyLeuArgAlaLeu   310
102499 AAGAGAGTGCTCGGGAGTAGAAAAGGGCTTTA 102528

```
seq_name: gb_ro:AY073444

seq_documentation_block:
LOCUS       AY073444               933 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR261-6 gene, complete cds.
ACCESSION   AY073444
VERSION     AY073444.1  GI:18480185
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)
  PUBMED    11802173
REFERENCE   2  (bases 1 to 933)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..933
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="6"
     mRNA            <1..>933
                     /product="olfactory receptor MOR261-6"
     CDS             1..933
                     /codon_start=1
                     /evidence=not_experimental
                     /product="olfactory receptor MOR261-6"
                     /protein_id="AAL61107.1"
                     /db_xref="GI:18480186"
                     /translation="MGNNMTLIEFILLGFPLSPRMQMLLFALFSLFYAFTLLGNGTI
                     VGLICLDSRLHTEPMYFFLSHLAIVDIAYACNTVPQMLVNLLDPVKPISYAGCMTQTFL
                     FLTFAITECLLLVVMSYDRYVAICHPLRYSAIMSWRVCSTMAVTSWIIGVLLSLIHLV
                     LLPLPFCVSQKVNHFFCEITALLKLACADTHLNETMVLAGAVSVLVGPFSSIVVSYA
                     CILGAILKIQSEEGQRKAFSTCSSHLCVVGLFYGTAIVMVGPRHGSPKEQKKYLLLF
                     HSLFNPMLNPLIYSLRNSDVKKNFLKRVLRTQRAL"
BASE COUNT      249 a    200 g    285 t
ORIGIN
                199 a    249 c    200 g    285 t

alignment_scores:
     Quality: 1330.00          Length:    310
     Ratio:   4.634            Gaps:      0
Percent Similarity: 92.581   Percent Identity: 81.935

alignment_block:
US-09-898-586-24 x AY073444

Align seg 1/1  to: AY073444  from: 1  to: 933
```

```
  1 MetGlyAsnAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
    |||||||||::||||::||:|||
  1 ATGGGAAACAATATGACATTAATCACAGAGTTCATCCTCCTGGGATTCC  50

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuProPheSerLeuPheT  34
    |:::||||||::|||||||||||:||||||:||:|||||||||||||||||||||
 51 TCTCAGCCCGAGGATGCAGATGCCTCCTTGCCCTTCTTCGCCCTTCTCTTCT  100

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu  50
    |||||||||||||:||||||||||||||||||||||||||||||||||||
101 ATGCCTTCACCCTGCTGGGGAAATGGAACATCGTGGGGCTGACTCATCCTG 150

 51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa  67
    ||||||||||||||||||||::||:::|||||||||||||::||||||||
151 GACTCCAGACTCCACACTCCCATGTACTTCTTCCTGTCCTCATGGGCCAT  200

 67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL  84
    |||||||||||||||||::|||||||||:::|||||||::|||||||::|
200 GACTCCAGACTCCACACTCCCATGTACTTCTTCCTGTCCTCATGGGCCAT  84
```

(Additional nucleotide/protein alignment blocks continue on the left column.)

seq_name: gb_pr:AC091768

seq_documentation_block:

```
LOCUS       AC091768        104955 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-703N5 from 7, complete sequence.
ACCESSION   AC091768
VERSION     AC091768.4  GI:15145624
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 104955)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 104955)
  AUTHORS   Harris,A., Haakenson,W. and Spalding,L.
  TITLE     The sequence of Homo sapiens BAC clone RP11-703N5
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 104955)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 104955)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 104955)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Aug 9, 2001 this sequence version replaced gi:14550325.
            --------------------  Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            --------------------  Summary Statistics
            Center project name: H_NH0703N05
            -----------------------------------
            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center.  For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7/, or send
            mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP4-669B10, 2000 bp overlap; the
            clone sequenced to the right is RP4-798C17, 2000 bp overlap.
            Actual start of this clone is at base position 12347 of
            RP4-669B10; actual end is at base position 64342 of RP11-798C17.
```

```
201 :;||||||:::||||||||:::||||||||:::::||||||||||||
    euLeuHisProAlaLysProSerPheAlaGlyArgMetGlnThr      250
    CGTTGACATTGCCTATGCCTGCAACACAGTGCCCAGATGCTGGTGAACC

 84 |||||
    euLeuHisProAlaLysProSerPheAlaGlyArgMetGlnThr      100
251 TTTTAGAATCAGTGAAGCCAGCCATTTCCTATGCTGGTTGTATGACACAGACT   300

101 |||||||||::|||||||||||||||||||||||::|||||||||||
    PheLeuPheSerThrPheAlaValThrGlyCysLeuLeuLeuValValMe  117
301 TTTCTTTTTTGCACATTTGCTATCACAGAATGCCTTCTCCTAGTGGTGAT      350

117 |||||||:::::||||
    tSerTyrAspLeuTyrValValAlaIleIleCysHisHisProLeuArgTyrLeuAlaI  134
351 GTCCTATGCTGCACGGTATGTTGCAAATCTGTCACCACCTCCGTTACTCTGCCA      400

134 |:::|||||||||||||||
    leMetThrTrpArgValCysIleThrLeuThrLeuAlaValThrSerTrpThrThr  150
401 TCATGACGTGGAGAGTTTGCACCATTGACATTACTCTTGCAGTGACTTCCTGGAATATT  450

151 |||||||||||
    GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuLeuProLeuProProh   167
451 GGTGTTCTGCTGTCTTTGATCCACCTCGTGTTACTTTTACCCTTACCCTTT      500

167 ::|||||||||||||||
    eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuLeuAlaLeuVaII   184
501 CTGTGTGTCTCAGAAAGTGAACCACTTTTTCTGTGAAATTACAGCTATCC      550

184 :::||||||||||||||
    euLysLeuAlaCysAlaLeuAspPhrHisIleAsnGluAsnMetValLeuAla   200
551 TCAAACTTGCCTGTGCAGACACACACCTCAATGAGAACATGGTCCTGGCT      600

201 ||||||||||||||||
    GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy   217
601 GGGGCAGTGTCTGGTCTTGTGGGGCCATTCTCCTCAAATTGTGGTCTCTTA      650

217 |||||||
    rMetCysIleLeuCysAlaIleIleLeuGlnSerArgGluValGlnGln   234
651 TGCATGTATCCTTGGTATCCTGAAGATCCAGTCTTTTAACCCTATGC      700

234 |||||||||
    rgLysAlaPheCysThrCysPheSerSerHisLeuCysValIleGlyLeuPhe   250
701 GGAAAGCCTTTTGCCACCTGCTTCCCCACCCTGTGTGTGGATAGGGATTGTTT      750

251 |||::||||||||||
    TyrGlyThrAlaIleIleMettTyrValGlyProArgTyrGlyAsnProLy   267
751 TATGGTACCGCCATTGTCATGTACGTTGGACCTCGTTATGGCAATACATTG      800

267 |||||||||||||||
    sGluGlnLysTyrLeuLeuPheHisSerLeuPheAsnProMetL   284
801 GGAGCAGAAGAAATATCTTTGCTGTTTCACAGTCTTTTTAACCCTATGC      850

284 |||||||||||
    euAsnProLeuIleCysSerLeuArgAsnSerGluIleLysSerAsnThrLeu   300
851 TCAACCCCTTGATCTATAGTCTCAGGAATTCTGACCTGAGAAATACATTG      900

301 ||||||||||
    LysArgValLeuGlyValGluArgAlaLeu   310
901 AAAAGAGTGCTGAGGAGACACAGAGAGCTTTG   930
```

```
                RP11-703N5 contains a single plasmid region from 4530 to 4616 and
                4668 to 4721.

FEATURES        Location/Qualifiers
source          1..104955
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="7"
                /map="7"
                /clone="RP11-703N5"
                /clone_lib="RPCI-11"
repeat_region   1..162
                /rpt_family="L1"
repeat_region   147..169
                /rpt_family="AT_rich"
repeat_region   928..1614
                /rpt_family="L2"
repeat_region   2051..2204
                /rpt_family="L2"
repeat_region   2649..2713
                /rpt_family="L2"
repeat_region   2651..2710
                /rpt_family="ERVL"
repeat_region   4071..4108
                /rpt_family="(CA)n"
repeat_region   4472..4511
                /rpt_family="(TCCA)n"
misc_feature    5679..5892
                /note="similar to Homo sapiens EST AA298029
                (NID:g1950363)"
repeat_region   7802..7825
                /rpt_family="AT_rich"
repeat_region   8020..8173
                /rpt_family="MIR"
repeat_region   9639..9865
                /rpt_family="CR1"
repeat_region   10327..10490
                /rpt_family="Alu"
repeat_region   10476..10498
                /rpt_family="Alu"
repeat_region   11042..11088
                /rpt_family="(CAAA)n"
repeat_region   11387..11406
                /rpt_family="GA-rich"
repeat_region   11947..11268
                /rpt_family="(T)n"
repeat_region   12489..13017
                /rpt_family="L1"
misc_feature    13193..13485
                /note="match to EST AI478847 (NID:g4372015) tm24e12.x1"
repeat_region   12974..13192
                /rpt_family="L1"
misc_feature    13147..13148
                /note="match to EST AI478847 (NID:g4372015) tm24e12.x1"
repeat_region   13386..13668
                /rpt_family="Alu"
repeat_region   13669..14177
                /rpt_family="MaLR"
repeat_region   14177..15248
                /rpt_family="MaLR"
repeat_region   15249..16333
                /rpt_family="MaLR"
repeat_region   16334..16622
                /rpt_family="ERVK"
repeat_region   16628..16988
                /rpt_family="MaLR"
repeat_region   16989..17116
                /rpt_family="MaLR"
repeat_region   17119..17273
                /rpt_family="MaLR"
repeat_region   17350..18588
                /rpt_family="L1"
repeat_region   18350..18588
                /rpt_family="ERVL"

repeat_region   16658..18733
                /rpt_family="ERVL"
repeat_region   19319..19982
                /rpt_family="L1"
repeat_region   20126..20161
                /rpt_family="AT_rich"
repeat_region   20389..21008
                /rpt_family="L1"
repeat_region   21009..23092
                /rpt_family="L1"
repeat_region   22685..22728
                /rpt_family="L1"
repeat_region   23114..23443
                /rpt_family="(TG)n"
repeat_region   23523..23592
                /rpt_family="L1"
repeat_region   25402..26143
                /rpt_family="AT_rich"
repeat_region   26433..26513
                /rpt_family="L1"
misc_feature    27066..27227
                /rpt_family="CR1"
repeat_region   27289..27962
                /rpt_family="MER1_type"
repeat_region   27954..28022
                /rpt_family="L1"
misc_feature    28774..29205
                /rpt_family="AT_rich"
misc_feature    28774..29234
                /note="similar to Homo sapiens EST BE179276
                (NID:g8658428)"
misc_feature    28774..29234
                /note="similar to Homo sapiens EST BF116115
                (NID:g10985591)"
misc_feature    28802..29204
                /note="similar to Homo sapiens EST H62441 (NID:g1015273)
                yr79d08.r1"
misc_feature    28816..29238
                /note="similar to Homo sapiens EST BE856570
                (NID:g10369724)"
misc_feature    28816..29238
                /note="similar to Homo sapiens EST AI079550 (NID:g3415801)
                oz04f08.x1"
misc_feature    28816..29238
                /note="similar to Homo sapiens EST AI148854 (NID:g3677323)"

alignment_scores:
    Quality:  1265.00          Length:    308
    Ratio:       4.600         Gaps:        0
    Percent Similarity: 89.286 Percent Identity: 80.195

alignment_block:
US-09-898-586-24 x AC091768

Align seq 1/1  to: AC091768    from: 1  to: 104955

    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
      ||||| ||||| |||||||||||||||||||||||||||||||||||||||
28794 ATGGGGGGGAAAATACAGACTTCCATCACAGAGTTCCTCCTACTGGGATTTCC 28843

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34
      |:::||||||||||||||||||||||||||||||||||||||||||||||
28844 CATTGGCCCAAGGAATTCAGAAGTGCTCCTCCTTGGGCTCTTCTCCCTGTTCT 28893

   34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuLeuSerLeu 50
      ||:::::||| |||||||||||||||||||||||||||||||||||||/
28894 ACATCTTCATTCTGTTGGGAACGGGACAAATCCTGGGGCTCATCTCACTG 28943

   34 AspSerArgIleAlaProMetArgTyrPhePheLeuSerHisLeuAlaVa 67
      |:::||||||||||||||||||:|||||||||||||||||||||||||||
28944 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCTCACACCTGGCGGT 28993

   51 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
      |||||||||||||||||:::|||||||||||||||||||||||||||||
28994 CGTCGACATCGCCTGTGCTTGCAGCACGGTGCCCCAGAA?GCTGGTGAACC 29043
```

```
PUBMED      11802173
REFERENCE   2 (bases 1 to 933)
  AUTHORS   Adams,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..933
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="6"
     mRNA            <1..>933
                     /product="olfactory receptor MOR261-1"
     CDS             1..933
                     /codon_start=1
                     /evidence=not_experimental
                     /product="olfactory receptor MOR261-1"
                     /db_xref="GI:18479498"
                     /protein_id="AAL60763.1"
                     /translation="MGGNQTLITQFILLGFPLSPRMQMLLFALFSLFYAFTLLGNGTI
                     LGLICLDSRLHTPMYFFLSHLAIVDIAYACNTVPQMLVNLMDPAKPISFAGCMTQFFL
                     FLTFAHTECLLLVMSYDRYVAICHPLRYTAIMSWRVCVILVLTISWLLGVLLALVHLV
                     LILPLPFCCSQKVNHFFCEIIAVLKLACSDTRINELMWVIAGAVSVLVGPFSSIVVSYA
                     HILCAILKIKSQQGRQKAFSTCSSHLCVVGLFYGTAIVMYIGPQHGKSNEQKKYLLLF
                     HSLFNPMLNPLIYSLRNKEVKSALKRTLLKEDTS"
BASE COUNT      201 a    248 c    195 g    289 t
ORIGIN

alignment_scores:
        Quality: 1246.00              Length:    304
        Ratio:   4.482                Gaps:        0
        Percent Similarity: 91.447    Percent Identity: 77.961

alignment_block:
US-09-898-586-24 x AY073100

Align seq 1/1    to: AY073100    from: 1    to: 933

1    MetGlyAspAsnIleThrSerIleThrGlnPheLeuLeuLeuGlyPhePr 17
     |||||| ||| |||||||::||||||||||||||||||||
1    ATGGGGGGAAATCAGACCTTGAATCACAGAGTTCAATCCTCCCTGGGATTCC 50

17   oValGlyProArgIleGlnMetLeuLeuPheAlaLeuPheSerLeuPheT 34
     |::::||||||||||||||||||:||||||||||||||
51   TCTCCAGCCCGAGGATGCAGATGCTCCTCTTCTTTGCCCTTTTCTCCCTGTTTCT 100

34   yrAlaPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
     |||||||||||||||||||||||||||||||:||||
101  ATGCCTTCACCCTGCTCGGGAAATGGGACCATCCTGGGACTTATCTGCCTG 150

51   AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
     |||||||||::  |||||||||||||||||||||||:
151  GACTCCAGACTCCACACTCCCATGTACTTCTTTCTGTCTCACCTGGCCAT 200

67   lValAspIleAlaTyrAlaCysAsnThrValProGlnMetLeuValAsnL 84
     :||||||||||||:||||||||||||||||||||
201  CGTTGACATTGCCTATGCCTGCAACACAGTGCCCCAGATGCTGGTAAACC 250

84   euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
     |:::||||||||||||||||||||||||||||||
251  TCATGGATCCAGCCAAGCCCATCTCTTTCGCTGGCCGTGGCATGACACAGAC 300

101  PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
     ||||||||||||||||||||||||||||||||||||||||||
301  TTTCTCTTTTTCGACGTTTGCACACACGGAAGTCTCCTACTCTGGTGGTGAT 350

117  tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
     |||||||||||||||||||||||||||||||||||||| ||
351  GTCCTATGATCGATATGTGAGCCATCCGCCACCACTCCGATATACTGCTA 400

134  leMetThrThrArgValCysIleThrLeuAlaLavalThrSerTrpThrThr 150
     |||||::::|||||||||||||||||||||||||
```

```
seq_documentation_block:
LOCUS       AY073100      933 bp    DNA     linear   ROD 04-FEB-2002
DEFINITION  Mus musculus olfactory receptor MOR261-1 gene, complete cds.
ACCESSION   AY073100
VERSION     AY073100.1  GI:18479497
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 933)
  AUTHORS   Zhang,X. and Firestein,S.
  TITLE     The olfactory receptor gene superfamily of the mouse
  JOURNAL   Nat. Neurosci. 5 (2), 124-133 (2002)

seq_name: gb.ro:AY073100

84    euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
      |||||:|||||||||||||||||||||||    |||
29044 TCCTGCACCAGCCAAGCCCATCTCCTTTGCTGGCCGTGCCATGACCCAGATG 29093

101   PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
      ||||||||||||||||||||||||||||:||||||||||:|:|
29094 TTTCTGTTTTCTAGTTTGCAATACAGAAGTCTCCTCTGGTGGTGAT 29143

117   tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
      |||||||||:|||||:||:|||||||||||||||||  :::|
29144 GTCCTATGATCGGTACGTGGGCCATGCCACCCTCTCCGATATTCTACCA 29193

134   leMetThrThrArgValCysIleThrLeuAlaLavalThrSerTrpThrThr 150
      |||||||||||||||||||||||||||||||||||
29194 TCATGACCACTGGAAAGTCTGCATCACTTTGGCATTGACTTCCTGGATTTTA 29243

151   GlyValLeuLeuSerLeuHisLeuValLeuLeuProLeuProPh 167
      ||||||||||||||||:||||||||||||||||||
29244 GGAGTCTTATTGGCCCTTGTCCATCTAGTGTTACTACCACACTGTCCTT 29293

167   eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184
      ||:::|||||||||||||||||||||||||
29294 CTGTGGACCCCAGAAACTTAATCACTTTTCTGTGAAATTATGGCTGTTC 29343

184   euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
      |||||||||||||||||||||||||||||||||
29344 TCAAACTTGCCTGTCGGGATACCACCATTAAATGAGGTAATGGTTTTGGCA 29393

201   GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
      :|||||  :::||||||||||||||||||
29394 GGGGCAGTGCTCGTGGTGGGGAGCCTTCTTCTTCTCCACTGTAATATCTTA 29443

217   rMetCysIleLeuLeuCysAlaIleLeuGlnIleLeuGlnSerArgValGlnA 234
      |||:||||||||||||||||||||||||||::|
29444 TGTTCATTATTCTATGTGCCATTCTAAAGATCCAGTCAGGAGAGGGGTGCC 29493

234   rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
      :||||||||||||||||||||||||||||:::
29494 AGAAAGGCCTTCTCCACTCTGCTCCTCCACCTGTGTGGTTGGGACTCTTT 29543

251   TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267
      ||||||||||||||:||||||||||||||
29544 TATGGCACAGCCATCATCATGTATGTTGGAGCCCAGATATGAGAGCCCCAA 29593

267   sGluGlnLysLysTyrrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
      :|||||||||||:|||::||||||||||||||||||
29594 GGAGCAGAAGAAATATCTCCGCTGTTTCACAGCCTCTTCAAATCCCATGC 29643

284   euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
      ||||||||||||||||||||||||||||||||
29644 TTAATCCCTAATTTAGTCTTAGGAACAAGGAAGTCCAAGGTACTCTA 29693

301   LysArgValLeuGlyValGluArg 308
      ||||||::|||
29694 AAGAGGATGCTTGAAAAGAAGAGA 29717
```

```
401 TCATGAGCTGGAGAGTTGTGTTATCCTGGTCGTGCTGACTTCCTGGATTTA 450
151 GlyValLeuSerLeuLeuIleHisIleuValLeuLeuLeuProLeuProh 167
451 GGAGTACTCCTGGCCCTGGCTCCANTCTAGTATTGTTGTTACCATTACCCTT 500
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVaIL 184
501 CTGCGGGCTCAACAAAAAGTGAAATCACTTTTTCTGTGAAATTATAAGCTGTTC 550
184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
551 TCAAACTTGCTTGTTCAGATACCCGCANTCAAATGAGTTAAATGGTTTTGGCT 600
201 GlyAlaAlaIleSerGlyLeuValGlyProLeuuSerThrIleValValSerTy 217
601 GGGGGTGTGTCTGTCGTGCGTGCCGACCATTTTGTTCAATTGTGAGTCTCTTA 650
217 rMetCysIleLeuCysSerAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
651 TGCTCATATATTCTGTGTGCCATCCTGAAGANTTAAGTCACAGCAGGGGCGGCC 700
234 rgLysAlaPheCysThrArgCysPheSerSerHisLeuCysValIleGlyLeuPhe 250
701 AGAAAAGCCTTCTCCACTTGCTCCTCCNTCCTGTGTTGTTGGGACTCTTT 750
251 TyrGlyThrArgAlaIleIleMetTyrvalGlyTyrProArgTyrGlyAsnProLy 267
751 TATGGTACACCAGCCATTATCATGTACANTGGGGCCTCAACACGGGAAGTCCAA 800
267 sGluGlnLeuLysThrTyrLeuLeuPheHisSerLeuPheAsnProMetL 284
801 TGAGCAGAGAAAATATCTCTTGCTATTCCATAGCCCTTTTAATCCTATGC 850
284 euAsnProLeuIleCysSerSerLeuArgAsnSerGluValLysAsnThrLeu 300
851 TCAACCCCTGATCTATAGTCTGAGGAACAACAAGAGGTCAAAAGTGCTCTG 900
301 LysArgValLeu 304
901 AAGAGGACACTC 912
```

seq_name: gb_pr:AC074386

seq_documentation_block:

```
LOCUS       AC074386    66610 bp    DNA    linear    PRI 07-NOV-2001
DEFINITION  Homo sapiens BAC clone RP11-464H1 from 7, complete sequence.
ACCESSION   AC074386
VERSION     AC074386.6    GI:13625507
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66610)
AUTHORS     Sulston,J.E. and Waterston,R.
TITLE       Toward a complete human genome sequence
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063792
REFERENCE   2  (bases 1 to 66610)
AUTHORS     Cordum,H. and Cotton,M.
TITLE       The sequence of Homo sapiens BAC clone RP11-464H1
JOURNAL     Unpublished
REFERENCE   3  (bases 1 to 66610)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (30-JUL-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 66610)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (15-APR-2001) Genome Sequencing Center, Washington
```

```
REFERENCE   5  (bases 1 to 66610)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (10-MAY-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 66610)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (07-NOV-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Apr 15, 2001 this sequence version replaced gi:13431207.
            --------------  Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            --------------  Summary Statistics
            Center project name: H_NH0464H01
            --------------
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7/, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An Improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)
VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP4-798C17, 200 bp overlap; the clone sequenced to the right is RP4-545C24, 200 bp overlap. Actual start of this clone is at base position 98638 of RP4-798C17; actual end is at base position 76592 of RP4-545C24.

There are polymorphic base pair differences in the overlap between the clone RP11-464H1 and RP4-545C24.

```
FEATURES             Location/Qualifiers
     source          1..66610
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP11-464H1"
                     /clone_lib="RPCI-11"
     misc_feature    98..259
                     /note="similar to EST AI603950 (NID:g4613112) ms16c11.y1"
```

```
misc_feature    98. .259   /note="similar to EST BE871418 (NID:g103201941)"
misc_feature    98. .259   /note="similar to EST BE383988 (NID:g9329353)"
misc_feature    98. .259   /note="similar to EST BE696685 (NID:g10083845)"
misc_feature   102. .259   /note="similar to EST BE383676 (NID:g9329139)"
misc_feature   114. .259   /note="similar to EST BE410379 (NID:g9346829)"
misc_feature   222. .259   /note="similar to EST BE410379 (NID:g9346829)"
repeat_region  297. .361   /note="similar to EST BG385901 (NID:g13279347)"
repeat_region  517. .646   /rpt_family="L2"
repeat_region  656. .951   /rpt_family="L1"
repeat_region 1149. .1451  /rpt_family="Alu"
repeat_region 1484. .1806  /rpt_family="Alu"
repeat_region 1743. .2035  /rpt_family="L2"
misc_feature  1942. .2074  /note="similar to EST AA078615 (NID:g18380096)"
repeat_region 2159. .2234  /rpt_family="7SK"
repeat_region 2159. .2234  /rpt_family="MIR"
repeat_region 2701. .2896  /rpt_family="MIR"
repeat_region 2908. .2949  /rpt_family="MIR"
repeat_region 3044. .3349  /rpt_family="MIR"
repeat_region 4064. .4319  /rpt_family="Alu"
misc_feature  4065. .4272  /note="similar to EST BF027677 (NID:g10735389)"
misc_feature  4066. .4319  /note="similar to EST D77287 (NID:g1597007)"
misc_feature  4066. .4319  /note="similar to EST BG385901 (NID:g13279347)"
misc_feature  4066. .4319  /note="similar to EST BE383676 (NID:g9329139)"
misc_feature  4066. .4319  /note="similar to EST BE871418 (NID:g103201941)"
misc_feature  4066. .4319  /note="similar to EST AI910192 (NID:g96500872)"
misc_feature  4066. .4265  /note="similar to EST BE383988 (NID:g9329353)"
misc_feature  4066. .4127  /note="similar to EST AI603950 (NID:g4613112) ms16c11.y1"
misc_feature  4066. .4319  /note="similar to EST BE696685 (NID:g10083845)"
misc_feature  4066. .4319  /note="similar to EST BE410379 (NID:g9346829)"
misc_feature  4164. .4319  /note="similar to EST AI910192 (NID:g96500872)"
misc_feature  4511. .4575  /note="similar to EST AI910192 (NID:g96500872)"
misc_feature  4511. .4575  /note="similar to EST BE383988 (NID:g9329353)"
misc_feature  4511. .4575  /note="similar to EST BE871418 (NID:g103201941)"
misc_feature  4511. .4575  /note="similar to EST BG385901 (NID:g13279347)"
misc_feature  4511. .4575  /note="similar to EST BE383676 (NID:g9329139)"
misc_feature  4511. .4575  /note="similar to EST BF027677 (NID:g10735389)"
misc_feature  4511. .4575  /note="similar to EST BE410379 (NID:g9346829)"
misc_feature  5009. .5066  /note="similar to EST BE383988 (NID:g9329353)"
misc_feature  5009. .5066  /note="similar to EST BE410379 (NID:g9346829)"
misc_feature  5009. .5066  /note="similar to EST BE383676 (NID:g9329139)"
misc_feature  5009. .5066  /note="similar to EST AI910192 (NID:g96500872)"
misc_feature  5009. .5066  /note="similar to EST BG385901 (NID:g13279347)"
misc_feature  5009. .5066  /note="similar to EST BG385901 (NID:g13279347)"
misc_feature  5009. .5066  /note="similar to EST BE871418 (NID:g103201941)"
misc_feature  5009. .5066  /note="similar to EST BF027677 (NID:g10735389)"
misc_feature  5598. .5674  /note="similar to EST BE871418 (NID:g103201941)"
misc_feature  5598. .5834  /note="similar to EST BF027677 (NID:g10735389)"
misc_feature  5598. .5729  /note="similar to EST BE383676 (NID:g9329139)"
misc_feature  5598. .5670  /note="similar to EST AI910192 (NID:g96500872)"
misc_feature  5598. .5670  /note="similar to EST BG385901 (NID:g13279347)"
misc_feature  5598. .5670  /note="similar to EST BE410379 (NID:g9346829)"
misc_feature  5696. .5670  /note="similar to EST BE383988 (NID:g9329353)"
misc_feature  5696. .6059  /note="similar to EST BF328558 (NID:g11299410)"
misc_feature  5718. .6272  /note="similar to EST AW784785 (NID:g9841561)"
misc_feature  5800. .5923
```

```
alignment_scores:
              Quality:  1138.00
              Ratio:       4.246
Percent Similarity:  88.158    Percent Identity:  72.039

alignment_block:
US-09-898-586-24 x AC074386

Align seq 1/1  to: AC074386   from: 1   to: 66610      Length:    304    Gaps:    0

      1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheP  17
        |||        ||| |||     :::|||||||||||||||||||||||
  21158 ATGGGTGAAAAATCAGACAATGGTCACAGAGAGTTCCTCCTACTGGGATTTCT 21207

     17 oValGlyProArgIleGlnMetLeuLeuPheGlyValPheSerSerPheT  34
        :::|||||:||||||||||||||||||||||||||||||||||||||||||
  21208 CCTGGGCCCAAGGATTCAGATACCTCCCTCTTTGGGGCTTCTTCTCCCGTTCT 21257

     34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu  50
        |||||||||||||||||||||||||||||||||||||||||||||||||||
  21258 ATGTCTTTCACCCTGCTGGGGAATGGGACCAACCTGGGGTCATCCTCACTG 21307

     51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa  67
        |||||||||:|||||||||||||:||||||||||||||||||||:|||||
  21308 GACTCCAGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGT 21357

     67 lValAspIleAlaTyrAlaCysSerAsnThrValProArgMetLeuValAsnL  84
        |||||:|||||||||||||||:|||||||||||||||::::|||||||||||
  21358 CGTCGACATCGCCTATGCCTGCAACACAGTGCCCCAGAATGCTGGTGAACC 21407

     84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
        ||||||||||||||||:|||||||||||||||||||||||||||||||||
  21408 TCCTGCATCCAGCCAACCCCATCTCCTTTGCTGGCCGCATGACATAGACC 21457

    101 PheLeuPheSerThrPheAlaValThrGluLeuLeuLeuLeuValValMe 117
        ||||||||||||||::|||||||||||||||||||||||||||||:::||
  21458 TTTCTCTTTTCTTTGAGTTTGCACATACTGAAGCCCTTCCTGTTGGTGCTGAT 21507

    117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
        |||||||||||||||||||||||||||||||||||||||||||||:|::
  21508 GTCCTAGATACGGTACCTGGCCATCTGCCACCCTCCGATATATTCATCA 21557

    134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
        ||||||||||||||:|:::||||||||||||||||||||::||||||||
  21558 TCATGACCTGGAGAAGTGTGCATCACTCTGGCCATCACTCCTGACAGT 21607
```

```
 151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167
     ||| |:|||:::|||||||||||||||||| ||:||||
21608 GGCTCCCCTCCTGGCTATGGTCCATTGTGAGCCTTCATCCTAAGACTGCCCTT 21657

 167 eCysArgProGlnLysIleTyrHisPheCysGluIleLeuAlaVaII 184
     |:|||||||||||||||||||||||||||||||:||
21658 TTGTGGGCCCTCGTGAAATCAACCACTTCTTGTGAAATCCTGTGTGTCC 21707

 184 euLysLeuAlaCysSerAlaAspThrHisIleAsnMetValLeuAla 200
     |:::||||||||||||||| :::::|||
21708 TCAGGCTGGCCTGTGCTCGATACCTGGCTCAACCAGGTGGTTCATCTTTGCA 21757

 201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
     ||:|/|| |/||||
21758 GCCCTGCCATGTTCATCCTGGGCAACCACTCTGCCTGGGTGCTCCCTA 21807

 217 rMetCysIleLeuCysAlaIleIleLeuGlnIleGlnSerArgGluValGlnA 234
     |||||||||||||||:|||||||| :::|
21808 CTCACACATCCTGGCGGGCCATCCTGAGGAATCCAGTCTGGGGAGGGCCGCA 21857

 234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
     |||||||||||||||||:::|||||||
21858 GAAAGGCCTTTCTCCACCTGCTCCTCCACCTGCCGTAGTGGGACTCTTC 21907

 251 TyrGlyThrAlaIleIleMetTyrValGlyProArgThryrGlyAsnProLy 267
     ::|||/||::||||||||||:||||||:
21908 TTTGGCAGCCGCCATCGTCATGTACATGGCCCCTAAGTCCCCCATCCTGA 21957

 267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
     :|||||:|||||||||||
21958 GGAGCAGCAGAAGGTCCTTTTTCTATTTTACAGTTCTTTCAACCCGATGC 22007

 284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
     ::|||||||||||||||| ||||||:::::|||
22008 TAAACCCCTGATTACAACCTGAGGAATGTAGAGGTCAAGGGTGCCCTG 22057

 301 LysArgValLeu 304
     :::||:::|||
22058 AGGAGGAGCACTG 22069
```

OM of: US-09-898-586-24 to: EST:*   out_format : pfs

Date: Aug 21, 2002 7:47 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xxlh
-Q=/cgn2_1/USPTO_spool/US09898586/runat_16082002_205735_20538/app_query.fasta_1.373
-DB=EST -OFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH-0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE-500
-MINLEN=0 -MAXLEN=200000000 -USER=US09898586_@CGN1_1_2668
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-898-586-24
Query length: 310
Database: EST:*
Database sequences: 13736207
Database length: -18414570560
Search time (sec): 1599.010000

score_list:

sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_htc:BC016940	+	1138.00	904.12	8.2e-97	2021	BC016940 Homo sapiens, Similar
gb_est1:AI148854	+	794.00	1333.22	5.2e-65	619	AI148854 qc70a02.x1 Soares_plac
gb_htc:AF327904	+	792.00	1328.62	9.3e-65	705	AF327904 Homo sapiens FKSG35 (F
gb_est2:BE856570	+	792.00	1328.32	1.0e-64	751	BE856570 7f64e08.x1 Soares_NSF_
gb_est1:AI339565	+	790.00	1329.32	8.5e-65	456	AI339565 qk67e07.x1 NCI_CGAP_Co
gb_est2:BE543056	+	762.00	1279.69	4.9e-62	580	BE543056 601068985F1 NIH_MGC_12
gb_est1:AI339573	–	754.00	1265.00	3.0e-61	443	AI339573 qk67f03.x1 NCI_CGAP_Co
gb_est2:BG984967	–	752.00	1265.30	3.1e-61	443	BG984967 IL5-CN0068-100401-434-
gb_est1:AI079550	+	750.00	1257.11	8.9e-61	738	AI079550 oz04f08.x1 Soares_feta
gb_gss:BH111304	+	742.50	1243.83	4.9e-60	788	BH111304 RPCI-24-367N6.TJ RPCI-
gb_est2:BF974217	+	687.00	1151.44	6.9e-55	671	BF974217 602243851F1 NIH_MGC_48
gb_est2:BF116115	+	678.00	1135.39	4.0e-54	732	BF116115 7n77g04.x1 NCI_CGAP_Ov
gb_est1:AI340119	+	643.00	1082.66	4.7e-51	367	AI340119 qk64f08.x1 NCI_CGAP_Co
gb_est1:AI285267	+	602.00	1013.45	3.3e-47	361	AI285267 qg138c08.x1 NCI_CGAP_Co
gb_est2:H62441	+	593.00	996.47	2.9e-46	426	H62441 yr79d08.r1 Soares_fetal
gb_est2:H62445	+	584.00	981.16	2.1e-45	439	H62445 yr79e08.r1 Soares_fetal
gb_est1:BE179276	+	581.00	971.84	6.9e-45	690	BE179276 RC1-HT0615-290300-021-
gb_est1:BE179276	+	581.00	971.84	6.9e-45	1394	AK017036 Mus musculus adult ma
gb_htc:AK017036	+	581.00	966.24	1.6e-44	1394	AK017036 Mus musculus adult ma
gb_htc:AK016560	+	580.00	955.16	2.2e-44	3063	AK016560 Mus musculus adult ma
gb_est1:AW299289	+	560.00	939.35	4.5e-43	499	AW299289 xs38f05.x1 NCI_CGAP_Ki
gb_est1:BE174029	–	554.00	930.05	1.4e-42	432	BE174029 QV1-HT0572-280200-089-
gb_est2:BE182734	–	522.50	879.73	9.4e-40	332	BE182734 RC3-HT0650-130400-011-
gb_est2:BF116115	–	517.50	857.06	4.8e-39	520	HS_5346_B2_F04_T7 RPCI
gb_gss:AQ600752	+	509.00	854.22	2.5e-38	441	AQ600752 CM3-MTO292-120101-616-
gb_est2:BIO20419	–	498.00	830.48	5.2e-37	762	BIO20419 RST12467 Athersys RAGE
gb_est2:BG193339	–	489.50	815.04	5.0e-36	853	BG193339 RST12467 Athersys RAGE
gb_gss:BH331857	–	474.00	794.84	5.0e-35	449	BH331857 CH230-46N15.TV CHOR1-2
gb_est2:BG197640	+	466.50	776.78	5.1e-34	796	BG197640 yp01h11.r1 Soares breast
gb_est1:BG197640	+	466.00	769.98	1.2e-33	1501	BG197640 RST17016 Athersys RAGE
gb_htc:AK016338	+	464.00	773.41	7.8e-34	726	AK225734 Mus musculus adult ma
gb_gss:AZ225734	+	464.00	773.46	7.8e-34	797	AZ225734 RPCI-23-16619.TJ RPCI-
gb_gss:AZ909618	+	455.50	758.16	5.5e-33	797	AZ909618 RPCI-24-222G18.TJ RPCI
gb_gss:AZ969227	+	455.00	759.34	4.8e-33	642	AZ969227 2M024LJ24R Mouse 10kb
gb_gss:BH078380	+	446.00	743.91	3.4e-32	656	BH078380 RPCI-24-349E9.TV RPCI-
gb_gss:BH078380	–	446.00	742.78	4.0e-32	740	AZ607393 1M0429N15R Mouse 10kb
gb_gss:AZ607393	–	442.50	737.49	7.8e-32	720	AZ607393 1M0429N15R Mouse 10kb
gb_gss:AZ090606	–	442.00	735.65	9.9e-32	842	AZ090606 RPCI-23-27B20.TV RPCI
gb_gss:AZ725747	–	442.00	737.81	9.9e-32	616	AZ725747 RPCI-24-98I12.TVB RPCI
gb_gss:AZ103967	+	438.50	731.81	7.5e-32	611	AZ103967 RPCI-23-33G18.TV RPCI
gb_gss:AZ353949	+	437.00	730.61	1.8e-31	534	AZ353949 1M0093C01F Mouse 10kb
gb_gss:AZ664867	+	437.00	730.61	1.8e-31	534	AZ664867 1M0545E05R Mouse 10kb
gb_gss:AG167722	+	434.00	723.29	4.8e-31	679	AG167722 Pan troglodytes DNA, c

seq_name: gb_htc:BC016940

seq_documentation_block:

LOCUS	BC016940			2021 bp	mRNA linear HTC 09-NOV-2001
DEFINITION	Homo sapiens, Similar to olfactory receptor, family 2, subfamily A,
	member 4, clone IMAGE:4424116, mRNA.
ACCESSION	BC016940
VERSION	BC016940.1 GI:16877381
KEYWORDS	HTC.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 2021)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (05-NOV-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgapbs-r@mail.nih.gov
	Tissue Procurement: ATCC
	cDNA Library Preparation: Life Technologies, Inc.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome
	Center, Stanford University School of Medicine, Stanford, CA 94305
	Web site: http://www-shgc.stanford.edu
	Contact: (Dickson, Mark) mcd@paxil.stanford.edu
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
	R. M.
	Clone distribution: MGC clone distribution information can be found
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
	Series: IRAK Plate: 27 Row: 1 Column: 22
	This clone was selected for full length sequencing because it
	passed the following selection criteria: Hexamer frequency ORF
	analysis, GenomeScan gene prediction, Similarity but not identity
	to protein
	This clone has the following problem: frame shifted.
FEATURES	Location/Qualifiers
source	1..2021
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:4424116"
	/tissue_type="Kidney, hypernephroma"
	/clone_lib="NIH_MGC_89"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"
BASE COUNT	470 a 546 c 427 g 578 t
ORIGIN

alignment_scores:
               Quality: 1138.00   Length: 304
               Ratio: 4.246   Gaps: 0
   Percent Similarity: 88.158   Percent Identity: 72.039

alignment_block:
US-09-898-586-24 x BC016940

Align seq 1/1 to: BC016940 from: 1 to: 2021

US-09-898-586-24 1/1 to: BC016940 from: 1 to: 2021

    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr   17
      ||| ||| ||| ||| |||:||||||||||||||||||||||||||||
  429 ATGGTGAAAAATCAGACAAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 478

sequence documentation:

	Strd					
gb_gss:AZ913406	–	432.50	719.28	8.1e-31	794	AZ913406 RPCI-24-165C12.TV R
gb_gss:AZ019257	–	432.00	720.62	6.8e-31	629	AZ019257 RPCI-23-266H21.TV R
gb_gss:BH299965	+	427.00	710.79	2.4e-30	728	BH299965 CH230-38K1.TJ CHORI
gb_gss:BH359995	+	427.00	710.43	2.5e-30	756	BH359995 CH230-12G16.TJ CHOR
gb_gss:AZ765752	+	425.00	708.73	3.1e-30	632	AZ765752 1M0562E19R Mouse 10

seq_name: gb_est1:AI148854
seq_documentation_block:
LOCUS   AI148854   619 bp   mRNA   linear   EST 28-OCT-1998
DEFINITION qc70a02.x1 Soares_placenta_8to9weeks_2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1714922 3', similar to SW:OLF7_MOUSE P34984 OLFACTORY
            RECEPTOR-LIKE PROTEIN K7 ;, mRNA sequence.
ACCESSION   AI148854
VERSION   AI148854.1 GI:3677323
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 619)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1098   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 432.
FEATURES   Location/Qualifiers
     source   1..619
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1714922"
                     /clone_lib="Soares_placenta_8to9weeks_2NbHP8to9W"
                     /dev_stage="two placentae: one from 8 weeks and another
                     from 9 weeks post conception"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGACGGCGGCCGCGATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo."
BASE COUNT   112 a   191 c   131 g   183 t   2 others
ORIGIN

alignment_scores:
                Quality: 794.00   Length: 198
                  Ratio: 4.511   Gaps: 0
     Percent Similarity: 88.889   Percent Identity: 76.768

alignment_block:
US-09-898-586-24 x AI148854

Align seg 1/1 to: AI148854   from: 1   to: 619

    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
      ||| ||| ||
   26 ATGGTGAAAAATCAGACAAATGGTCACAGAGTTCCTCTACTGGGAATTCT 75

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
      ::::|||||||||||||||||||||||||||||||||||||||||||||
   76 CCTGGGGCCCAAGGATTCAGATGCTCCCCTCTTTGGGCTTCTTCCCTGTTCT 125

   34 yrValPheThrLeuLeuGlyLeuAsnGlyThrIleLeuGlyLeuIleSerLeu 50
      :|||||||||||||||||||||||||||||||||||||||||||||
  126 ATGGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGCTCATCTCACTG 175

   51 AspSerArgLeuHisIleAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
      |||||||||||||||||||||||||||||||||||||||||||||||||||
  176 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCTCACACCTGGCCGT 225

seq_name: gb_htc:AF327904

LOCUS　　　　　AF327904　　　　705 bp　　　mRNA　　　linear　　　HTC 13-JUL-2001
DEFINITION　Homo sapiens FKSG35 (FKSG35) mRNA, complete cds.
ACCESSION　　AF327904
VERSION　　　AF327904.1　　GI:12276181
KEYWORDS　　　HTC.
SOURCE　　　　human.
ORGANISM　　　Homo sapiens
　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　　1　(bases 1 to 705)
AUTHORS　　　Wang,Y.-G. and Gong,L.
TITLE　　　　Cloning of FKSG35, a novel gene located on human chromosome 7
JOURNAL　　　Unpublished
REFERENCE　　2　(bases 1 to 705)
AUTHORS　　　Wang,Y.-G.
TITLE　　　　Direct Submission
JOURNAL　　　Submitted (13-DEC-2000) Beijing Fengkesheng Function Gene
　　　Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,
　　　Beijing 100050, P.R. China
FEATURES　　　　　　　　　Location/Qualifiers
　　　source　　　　　　1..705
　　　　　　　　　　/organism="Homo sapiens"
　　　　　　　　　　/db_xref="taxon:9606"
　　　　　　　　　　/chromosome="7"
　　　　　　　　　　/map="7q33-q35"
　　　　　　　　　　/tissue_type="liver"
　　　gene　　　　　　　1..705
　　　　　　　　　　/gene="FKSG35"
　　　CDS　　　　　　　133..429
　　　　　　　　　　/gene="FKSG35"
　　　　　　　　　　/codon_start=1
　　　　　　　　　　/product="FKSG35"
　　　　　　　　　　/protein_id="AAG50283.1"
　　　　　　　　　　/db_xref="GI:12276182"
　　　　　　　　　　/translation="MVKNQMVTEFLLLGFLLGPRIQMLLFGLFSLFYVFTLLGNGTI
　　　　　　　　　　LGLISLDSRLHPTMPYFFLSHLAVVNIAYACNTVPQMLVNLLHPAKPISFAGCMT"
BASE COUNT　　　156 a　　　201 c　　　132 g　　　216 t

ORIGIN
67　IValAspIleAlaThrTyrAlaCysAsnThrValProArgMetLeuValAsnI　84
226　CGTCAACGTCGCCTATGCCTGCAACACAGTGCCCAGATGCTGGTGAACC　275

84　euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr　100
276　TCCTGCAACCCAGCCAAGCCCAATCTCCTTTGCTGGCCTGCATGACAATAGACC　325

101　PheLeuPheSerThrPheAlaLeuValThrGluCysLeuLeuLeuValVaIMe　117
326　TTTCTCTTTTTTCGAGTTTTGCACATACTGAAATGCCTCCCGTTGGGTGCTGAT　375

117　tSerTyrArgAspLeuThrYrValAlaIleCysHisHisProLeuArgTyrLeuAlaI　134
376　GTCCTACGATCGGTACCTGGCCATCTGCCACCACCCTCCGATAATTCATCA　425

134　leMetThrTrpArgValCysIleThrLeuAlaIValThrSerTrpThrThr　150
426　TCATGACCTGNAAAGTGCATACACTCTGGCCATACATTCCTGACACATGT　475

151　GlyValLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProh　167
476　GGCTCCTCCTGGCTANGGNCCATGTGAGCCTCATCCTAAGACTCCCCTT　525

167　eCysArgProGlnLysIleTyrHisPhePheCysGluIleIleLeuAlaVaIL　184
526　TTGTGGGGCTCCGTGAAATCAACCACCACTTCTTCTGTGAAATCCTGGCACTG　575

184　euLysLeuValAlaCysAlaAspThrThrHisLeuAsnGluAsnMetVal 198
576　TCAGGCTGGGCTCTGCTGCATAACTGGCCTCAACCAGGTGGTCATC 619

alignment_block:
US-09-898-586-24 x AF327904

Align seq 1/1　to: AF327904　from: 1　to: 705

1　MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr　17
133　ATGGTGAAAAATCAGCACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT　182

17　oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT　34
183　CCTGGGCCCAAGGATTCAGAATGCTCCTCTTTTGGGCTCTTCTCCCTGTTCT　232

34　yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu　50
233　ATGTCTTCACCCTGCTGGGGAATGGGACCATATCCTGGGGCTCAATCTCACTG　282

51　AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa　67
283　GACTCCAGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGT　332

67　IValAspIleAlaThrTyrAlaCysAsnThrValProArgMetLeuValAsnL　84
333　CGTCAACATCGCCTATGCCTGCAACACAGTGCCCAGATGCTGGTGAACC　382

84　euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr　100
383　TCCTGCATCCAGCCAAGCCCAATCTCCTTTGCTGGCCTGCATGACAATAGACC　432

101　PheLeuPheSerThrPheAlaLeuValThrGluCysLeuLeuLeuValVaIMe　117
433　TTTCTCTTTTTTCGAGTTTTGCACATACTGAAATGCCTCCTGTTGGGTGCTGAT　482

117　tSerTyrArgAspLeuThrTyrValAlaIleCysHisHisProLeuArgTyrLeuAlaI　134
483　GTCCTACGATCGGTACCTGGCCATCTGCCACCACCCTCCGATAATTCATCA　532

134　leMetThrTrpArgValCysIleThrLeuAlaIValThrSerTrpThrThr　150
533　TCATGACCTGGAAAGTGCATACACTCTGGCCATACATTCCTGACACATGT　582

151　GlyValLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProh　167
583　GGCTCCTCCTGGCTATGGTCCATGTGAGCCTCATCCTAAGACTCCCCTT　632

167　eCysArgProGlnLysIleTyrHisPhePheCysGluIleIleLeuAlaVaIL　184
633　TTGTGGGGCTCCGTGAAATCAACCACCACTTCTTCTGTGAAATCCTGGCACTG　682

184　euLysLeuValAlaCysAlaAspThr　191
683　CTCAGGCTGGGCTCTGCTGCATAACTACC　705

alignment_scores:
　Quality:　792.00　　Length:　　191
　Ratio:　　4.632　　Gaps:　　　0
Percent Similarity:　89.529　Percent Identity: 79.581

seq_name: gb_est2:BE855670

seq_documentation_block:

LOCUS　　　BE855670　　　751 bp　　mRNA　　　linear　　EST 29-SEP-2000
DEFINITION　7f64e08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
　　　IMAGE:3299462 3' similar to TR:O95047 O95047 WUGSC:H_DJ0988G15.2
　　　similar to TR:O95047 O95047 WUGSC:H_DJ0988G15.2
PROTEIN　　　;, mRNA sequence.
ACCESSION　　BE855670
VERSION　　　BE855670.1　　GI:10369724
KEYWORDS　　　EST.
SOURCE　　　　human.
ORGANISM　　　Homo sapiens

Left column

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 751)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 449.
FEATURES             Location/Qualifiers
     source          1..751
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3399462"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     15407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320135-322823, 326280-326663 Soares NbHOT
                     pool 1: 723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT   170 a   206 c   137 g   238 t
ORIGIN
```

alignment_scores:
```
               Quality:   792.00          Length:      191
Percent Similarity: 89.529          Gaps:        0
               Ratio:    4.632          Percent Identity:  79.581
```

alignment_block:
US-09-898-586-24 x BE855570

Align seg 1/1 to: BE855570   from: 1 to: 751

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr  17
      ::||||||||||||||||||||||||||||||||||||||||||||||
179 ATGGTGAAAAATCAGAACAAGGTCACAGAGTTCCTCCTACTAGGATTTCT 228

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT  34
      ::|||||||||||||||||||||||||||||||||||||||||||||
229 CCTGGGCCCAAGGATTCAGATGCTCCTCTGTTTGGGCTCTTTTCTCTGTTTCT 278

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu  50
      |||||||||||||||||||||||||||||||||||||||||
279 ATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGTCCATCTCACTG 328

 51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa  67
      ||||||||||||||||||||||||||||||||||||||||||||||
329 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCTCACACCTGGCCGT 378

 67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL  84
      ||||||||||||||||||||||||||||||||||||||||||
379 CGTCAACATCGCCTAYGCCTGCAACACAGTGCCCCAGAGCTGGTGAACC 428

 84 euLeuHisProAlaLysProPheSerPheAlaGlyArgMetMetGlnThr 100
      ||| ||||||||||||||||||||||||||||||||||| ||| |||
429 TCCTGCATCCAGCCAAGCCATTCTCCTTTGCTCTGGCTGGCATGACATAGACC 478
```

Right column

```
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValVaLMe 117
      ||||||||| ||||||||||||||| ||||||||||||||::|||
479 TTTCTCTTTTTGAGTTTTGCACATACTGAATGCCTCCTCGTTGGTGCTGAT 528

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
      ||||||||| ||||||||||||||||||||||||||||::|
529 GYCCTACGATCGGTACGTGGCCCATCTGCCACCCCTCCGATATTTCATCA 578

134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
      |||||| |||||||||||||| |||:|||:||||||||||||||
579 TCATGACCTGGAAAGTCTGCATCACTCTTGCCGTCACCTCCTGGACATGT 628

151 GlyValLeuLeuSerIleLeuHisLeuValLeuLeuLeuProLeuProH 167
      ||||||||| |||||||||::|||:|||||||||||||||||||||
629 GGCTCCCTCCTGGCTATGGCGTCCATGTAGCCCTAAGACTGCCCTT 678

167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVaLl 184
      ||| |||||||||||||||||| |||||||||||:|||||
679 TTGTGGGCCTCGTGAAATCAACCACTCTTCTGTGAAAAACCTGTCTGTCC 728

184 euLysLeuAlaCysAlaAspThr 191
      ::||||||||||||::|||||||||
729 CTCAGCTGGCCTGTGCTGATACC 751
```

seq_name: gb_est1:AI339565

seq_documentation_block:
```
LOCUS       AI339565         456 bp    mRNA    linear   EST 13-FEB-1999
DEFINITION  qk67e07.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874052 3'
            similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
            ; mRNA sequence.
ACCESSION   AI339565
VERSION     AI339565.1  GI:4076492
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 456)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            Insert Length: 669   Std Error: 0.00
FEATURES             Location/Qualifiers
     source          1..456
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1874052"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector. Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo."
BASE COUNT    136 a    93 c    149 g    78 t
```

Left column

ORIGIN

alignment_scores:
```
    Quality:   790.00              Length:     151
    Ratio:     5.232               Gaps:         0
Percent Similarity: 100.000    Percent Identity: 100.000
```

alignment_block:
US-09-898-586-24 x AI339565/rev

Align seg 1/1 to reverse of: AI339565   from: 1 to: 456

```
 31 SerLeuPheTyrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLe  47
454 TCCCTGTTCTACGTCTTCACCCTGCTGGGAAACGGGACCATACTGGGGCT 405

 47 uileSerLeuAspSerArgLeuHisAlaProMetTyrPhePheLeuSerH  64
404 CAACTCACTGGGACTCCAGACGCACGCCCCCCATGTACTTCTTCCTCTCAC 355

 64 isLeuAlaValValAspIleAlaTyrAlaCysHisAsnThrValProArgMet  80
354 ACCTGGGCCGGTCGTCGACAATCGCCTACGCCTGCAACACGGTCCCCGGATG 305

 81 LeuValAsnLeuLeuHisProAlaLysProIleSerPheAlaGlyArgMe  97
304 CTGGTGAACCTCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGGCGCAT 255

 97 tMetGlnThrPheLeuPheSerThrPheAlaValThrGluCysLeuLeuL 114
254 GATGCAGACCTTTCTGTTTTCCACTTTTGCCGTCACAGAANGTCCCTCC 205

114 euValValMetSerTyrTyrAspLeuTyrValAlaIleIleCysHisProLeuArgMe 130
204 TGGTGGTGATGTCCTATACTTACGTGGCGCCATCTGCCACCCTCCGA 155

131 TyrIleLeuAlaIleLeuMetThrThrArgValCysIleThrLeuAlaLeuPhrSe 147
154 TATTTGGCCATCANGACCTGGACAGTGGAGAGTTCGCATCACCCTCGCGGTGACTTC 105

147 rTrpThrThrGlyValLeuLeuSerLeuGlnHisIleLeuValLeuLeuLeuP 164
104 CTGGACCACTGGAGTCCTTTTATCCTTGATTCTNCTTGTTGTTACTTCTAC 55

164 roLeuProPheCysArgProGlnLysIleTyrHisPhePheCysGluIle 180
 54 CTTTACCCTTCTCTAGGCCCCCAGAAAATTATCACTTTTTTTTGTGAAATC 5

181 Leu 181
  4 TTA 2
```

seq_name: gb_est2:BE543056

seq_documentation_block:

```
LOCUS       BE543056       580 bp    mRNA    linear   EST 09-AUG-2000
DEFINITION  601068985F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3455324 5',
            mRNA sequence.
ACCESSION   BE543056
VERSION     BE543056.1  GI:9771701
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
```

Right column

```
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8441 Row: i column: 21
            High quality sequence start: 3
            High quality sequence stop: 580.

FEATURES             Location/Qualifiers
     source          1..580
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3455324"
                     /clone_lib="NIH_MGC_12"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
BASE COUNT     96 a    189 c    125 g    170 t
ORIGIN
```

alignment_scores:
```
    Quality:   762.00              Length:     193
    Ratio:     4.379               Gaps:         0
Percent Similarity: 90.155     Percent Identity: 77.202
```

alignment_block:
US-09-898-586-24 x BE543056

Align seg 1/1 to: BE543056   from: 1 to: 580

```
 22 IleGlnMetLeuLeuProPheGlyLeuPheSerLeuPheThrTyrValPheThrIle 38
  3 ATTCAGATGCTCCTTCTTTTGGGCTTTCCTCCCTGTTCTTTATGTCTTCACCCT 52

 38 uLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSerArgLeuH 55
 53 GCTGGGAAATGGGACCATCCTGGGGCTCATCTCACTGGACTCCAGACTC 102

 55 isAlaProMetTyrPhePheLeuSerHisLeuAlaValValAspIleAla 71
103 ACACCCCCATGTACTTCTTCCTCTCACACCCTGCCGTCGTCGACAATCGCC 152

 72 TyrAlaCysHisAsnThrValProArgMetLeuValAsnLeuLeuHisProAl 88
153 TATGCCTGCCAACACAGTGCCCCAGATGCTGGTGAACCTCCTGCATCCAGC 202

 88 aLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeuPheSerT 105
203 CAAGCCCATCTCCTTTGCCGGGCGTGCGCATGATGACATAGACCTTCTCTTTTGA 252

105 hrPheAlaValThrGluCysLeuLeuLeuValValMetSerTyrTyrAspLeu 121
253 GTTTTGCACATACTGAATCGCTCCTCCTGTTGGTGCTGAGTCCTACGATCGG 302

122 TyrValAlaIleIleCysHisProLeuArgTyrIleLeuAlaIleLeuMetThrThrAr 138
303 TACCTGGCCATCTGCCACCCTCTCCGATAATTTCATCACTAACCTCGGAA 352

138 gValCysIleThrLeuAlaLeuAlaThrSerTrpThrThrGlyValLeuLeuS 155
353 AGTCTGCATCACNCTGGCATCACTTCCTGGACATGGGGCTCCCTCCTGG 402

155 erLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArgProGln 171
403 CTATGGTCCCATNGTGAGCCTTCATCCTAAGACTGCCCTTTGTGGGGCTCGT 452

172 LysIleTyrHisPhePheCysGluIleLeuLeuAlaValLeuLysLeuAlaCy 188
```

seq_name: gb_est1:AI339573

```
453 GAAATCAACCACTCTCTTCTGTGAAATCCTGTCGTCGTCTGTCCTTCAGGCGTGGCCTG
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
188 sAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAlaIleSerG 205
503 TGCTGATATACCTGGGCGTCAACCAAGGTGGTCATCTTTGCGCCTGCATGTTCAT 551
        |||||||||||||||||||||||||||||||||||||||||||||
205 lyLeuValGlyProLeuSerThrIleVal 214
        : : : :
552 C.CTGGTGGGACCACTCTGCCCTGTGTGCTG 580
```

seq_documentation_block:
| | |
LOCUS   AI339573   433 bp   mRNA   linear   EST 13-FEB-1999
DEFINITION   qk67f03.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874045 3'
                similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
                ; mRNA sequence.
ACCESSION   AI339573
VERSION   AI339573.1 GI:4076500
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 433)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
                Emmert-Buck, M.D., Ph.D.
                cDNA Library Preparation: M. Bento Soares, Ph.D.
                cDNA Library Arrayed by: Greg Lennon, Ph.D.
                DNA Sequencing by: Washington University Genome Sequencing Center
                Clone distribution: NCI-CGAP clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                www-bio.llnl.gov/bbrp/image/image.html
                Insert Length: 677   Std Error: 0.00
                Seq primer: -40UP from Gibco
                High quality sequence stop: 411.
FEATURES   Location/Qualifiers
     source   1..433
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1874045"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector. Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo."
BASE COUNT   127 a   91 c   139 g   76 t
ORIGIN
```

alignment_scores:
```
        Quality: 754.00   Length: 144
        Ratio: 5.236   Gaps: 0
Percent similarity: 100.000   Percent Identity: 100.000
```

alignment_block:
US-09-898-586-24 x AI339573/rev

Align seg 1/1  to reverse of: AI339573   from: 1 to: 433

```
38 LeuLeuGlyAsnGlyThrIleLeuIleSerLeuAspSerArgIle 54
```

---

seq_name: gb_est2:BG984967

```
433 CTGCTGGGAACGGGACGGACCAATACTGGGGCTCCATCCACTGGACTCCAGACT 384
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
54 uhisAlaProMetTyrPhePheLeuSerHisLeuValAlaValAspIleA 71
383 GCACGCCCCATGTACTTCTTCCTCCACAACCTGGCGGTGGTCGACATCG 334
        ||||||||||||||||||||||||||||||||||||||||||||||
333 CCTACGCCTGCAACACGCGTGCCCCGGAATGCTGGTTGGTGAACCTCCTGCATCCA 284
        ||||||||||||||||||||||||||||||||||||||||||||||
71 laTyrArgAlaCysAsnThrValProArgMetLeuValAsnLeuLeuHisPro 87
283 GCCAAGCCCATCTCCTTTGCGGGGCGGCATGAATGGCAGAACCTTTCTGTTTTC 234
        ||||||||||||||||||||||||||||||||||||||||||||||
88 AlaLysProIleSerPheAlaGlyArgMetGluGlnThrPheLeuPheSe 104
233 CACTTTTGCTGTCACAGAAATGTCTCCTCCCTCCTTCCGGTGGGTGGTGATCTCGTCATGAATC 184
        ||||||||||||||||||||||||||||||||||||||||||||||
104 rThrPheAlaValThrGluCysLeuLeuLeuLeuValValMetSerTyrAspL 121
183 TGTACGTGGCCATCTGCCACCCCCCGGATATTTGGGCCATCATGACCTGG 134
        ||||||||||||||||||||||||||||||||||||||||||||||
121 euTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMetThrTrp 137
154 uSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArgProG 171
        ||||||||||||||||||||||||||||||||||||||||||||||
133 AGAGTCTGCATCACCCCTCCGCGGTGACTTCCTGGACCACTGGAGTCCTTTT 84
83 ATCCTTGATTCATCCTTGGTTGTTACTTCTACCTTTACCCTTCTGTAGGCCC 34
        ||||||||||||||||||||||||||||||||||||||||||||||
154 uSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArgProG 171
171 lnLysIleTyrHisPhePheCysGluIleLeu 181
        ||||||||||||||||||||||
33 AGAAAATTTATCCACTTTTTTTTTTGTGAAATCTTA 2
```

seq_documentation_block:
```
LOCUS           BG984967              443 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION      IL5-CN0068-100401-434-d09 CN0068 Homo sapiens cDNA, mRNA sequence.
ACCESSION       BG984967
VERSION         BG984967.1  GI:14387702
KEYWORDS        EST.
SOURCE          human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1  (bases 1 to 443)
AUTHORS         Dias,Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
                Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
                Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
                Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
                ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
                Simpson,A.J.
TITLE           Shotgun sequencing of the human transcriptome with ORF expressed
                sequence tags
MEDLINE         20202663
JOURNAL         Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT         Contact: Simpson A.J.G.
                Ludwig Institute for Cancer Genetics
                Laboratory of Cancer Research
                Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                Brazil
                Tel: +55-11-27704922
                Fax: +55-11-2707001
                Email: asimpson@ludwig.org.br
                This sequence was derived from the FAPESP/LICR Human Cancer Genome
                Project. This entry can be seen in the following URL
                (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-CN0068-
                100401-434-d09&t3=2001-04-10&t4=1)
                Seq primer: puc 18 forward
                High quality sequence stop: 363.
FEATURES             Location/Qualifiers
```

source
1. .443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0068"
/dev_stage="Adult"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT   136 a   94 c   131 g   81 t   1 others
ORIGIN

alignment_scores:
Quality: 752.00
Ratio: 5.151
Percent Similarity: 99.320   Percent Identity: 98.639

Align seq 1/1 to reverse of: BG984967   from: 1 to: 443

alignment_block:
US-09-898-586-24 x BG984967/rev

Length: 147
Gaps: 0

```
70  IleAlaThrValAlaCysAsnThrValProArgMetLeuValAsnLeuLeuHi  86
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
443 ATCGCCTACGCCGTCGCTTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTGCA 394

86  sProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeuP  103
    ||||||||||||||||||||||||||||||||||||||||||||||||||
393 TCCAGCCAAGCCCATCTCCTTNCCGGGCCGCCCATGATCAGACCTTTCTGT 344

103 heSerThrPheAlaValThrGlyCysIleLeuLeuLeuValValMetSerTyr 119
    |||||||||||||||||||:|::|||||||||||||||::::||||||||||
343 TTTCCACTTTGCTGTCACAGAATGTCTCCCTGGTGGTGATGTCCTAT 294

120 AspLeuTyrValAlaIleIleCysHisProLeuArgTyrLeuAlaIleMetth 136
    |||||||||||||||||||||||||||||||||||||||||||||||||||
293 GATCTGTACGTGGCCATCTGCCACCCCCTCCGATAYATTTGGCCATCATGAC 244

136 rThrProArgValCysIleThrLeuAlaValThrSerTrpThrThrGlyValL 153
    ||||||||||||||||||||||||||||||||||||||||||||||||||
243 CTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTCGGACACTGAGTCC 194

153 euLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArg 169
    ||||||||||||||||||||||||||||||||||||||||||||||||||
193 TTTTTATCCTTGAYTCATCTTGYTGTTACTTCTTTACCCTTCTGTAGG 144

170 ProGlnLysIleTyrHisPhePheCysGluIleLeuLeuAlaValLeuLysSe 186
    |||||||||||||||||||||||||||||||||||||||||||||||||
143 CCCCAGAAAATTTATCACTTTTTTGTGAAANTCTGGCTGYTNCTCAAACT 94

186 uAlaCysSAlaAspThrHisIleAsnGlyAsnMetValLeuuAlaGlyVaLAl 203
    |||||||||||||||||||||||||||||||||||||||||||||||||
93  TGCCTGTGCAGATACCACACATCAAATAGAACAATGGTCTTGGCCGGAGCAA 44

203 leSerGlyLeuValGlyProLeuSerPheThrIleValValSer 216
    |||||||||||||||||||||||||||||||||||||||||||||
43  TTTTCTGGGCTGGGTGGGACCCTTGTCCACAATTGTAGTTTCA 3
```

seq_documentation_block:
seq_name: gb_est1:AI079550

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 738)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1896   Std Error: 0.00
            Seq primer: -40m13 fwd ET from Amersham
            High quality sequence stop: 447.

FEATURES   Location/Qualifiers
source   1. .738
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1674375"
                /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                /sex="male"
                /dev_stage="20 week-post conception fetus"
                /lab_host="DH10B (ampicillin resistant)"
                /note="Organ: Liver and Spleen; Vector: p77T3D (Pharmacia)
                with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                This is a subtracted version of the original Soares fetal
                liver spleen 1NFLS library. 1st strand cDNA was primed
                with a Pac I - Oligo(dT) primer [5'
                AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Pac I and cloned into the Pac I
                and Eco RI sites of the modified pT7T3 vector. Library
                went through one round of normalization. Library
                constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT   173 a   201 c   135 g   229 t
ORIGIN

alignment_scores:
Quality: 750.00
Ratio: 4.545
Percent Similarity: 89.189   Percent Identity: 78.919

Align seq 1/1 to: AI079550   from: 1 to: 738

alignment_block:
US-09-898-586-24 x AI079550

Length: 185
Gaps: 0

```
1   MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
184 ATGGTGAAAAATCAGACAAATGGTCACACAGAGTTCCTACTCACTGGATTTCT 233

17  oValGlyProArgIleGlnMetLeuuPheGlyLeuPheSerLeuPheF  34
    ::|||||||||||||||||||||||||||||||||||||||||||||||
234 CCTGGGCCCCAAGGATTCAGATGCTCCTCTTTGGGCGTCTTCTCCCGTGTTCT 283

34  yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu  50
    |||||||||||||||||||||||||||||||||||||||||||||||||
284 ATGTCTTCACCCTGCTGGGGAAYGGGACCAATCCTGGGGCTCAYCTCACTG 333

51  AspSerArgLeuHisAlaProMetThrTyrPhePheLeuSerHisLeuAlaVa 67
    |||||||||||||||||::|||||||||||||||||||||||||||||
334 GACTCCAGACTCCACACCCCCAYGTACTTCTTCCTCTCACACCTGGCCGT 383

67  lValAspIleuAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
    ||:||||||::|||||||||||||||||||||||||||||||||||||||
384 CGTCAACATCGCCTAYGCCTGCAACACAGTGCCCCAGAATGCTGGTGAACC 433

84  euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
    |||||||||||||||||||||||||||||||||||||||||||  |||
434 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGCCTGCATGACATAGACC 483
```

seq_name: gb_gss:BH111304

seq_documentation_block:
LOCUS   BH111304   788 bp   DNA   linear   GSS 19-JUL-2001
DEFINITION RPCI-24-367N6.TU RPCI-24 Mus musculus genomic clone RPCI-24-367N6,
            DNA sequence.
ACCESSION   BH111304
VERSION   BH111304.1 GI:14946319
KEYWORDS   GSS.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 788)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE   Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/ordering/ordering_frame.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_ends_intro.html
            Plate: 367 row: N column: 6
            Seq primer: SP6
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..788
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-367N6"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
                     RPCI-24 Mouse BAC library produced by Pieter de Jong. The
                     library was cloned in the pTARBAC1 cloning vector at the
                     BamHI sites using MboI partially digested male C57BL/6J
                     DNA."
BASE COUNT   168 a   217 c   162 g   241 t

ORIGIN

alignment_scores:
     Quality: 742.50   Length: 264
     Ratio: 3.486   Gaps: 5
     Percent Similarity: 80.682   Percent Identity: 57.955

alignment_block:
US-09-898-586-24 x BH111304

Align seq 1/1 to: BH111304

Align seq 1/1 to: BH111304   from: 1 to: 788

```
101 PheLeuPheSerThrThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
484 TTTCTCTTTTGAGTTTTGCACATACATGAATGACCTGCTCGTTGGTGCTGAT 533
    |||||||   ||||||||||||||||||||||||||||||||||||||||| |

117 tSerTyrAspLeuThrTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
534 GTCCTACGATACGGTACGTGGCCATCTGCCACCCTTCCCGATATTTCATCA 583
    |  |||  :    |:::|||||||||||||||||||||||:::|   | |

134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrhrHr 150
584 TCATGACCTGGAAAGTGTCGATCACTCTCGGCCATCACTTCCTGGACAHGT 633

151 GlyValLeuLeuSerLeuLeuLeuHisLeuValLeuLeuLeuProLeuProh 167
634 GGCTCCCTCCTGGCTATGGTCCCAGTGAGCCCTCATCCTAAGACTGGCCCT 683
    |||   |||||||||:|||||:||||:|||||||:||||     |||  ::

167 eCysArgProGlnLysIleThrTyrHisPhePheCysGluIleLeuAlaVal 184
684 TTTTGTGGCCTCGTGAAAATCAACCACTTTCCTCGTGAAAATCCTTGCTGT 733
    :  ||||  :||:||||||||||||||||||||  |||||| ||||||||||

184 euLys 185
734 TCAAG 738
     ||||
```

alignment_scores (detail):
```
34  TyrValPheThrIleLeuGluGlyAsnGlyThrIleLeuIleIleSerIle 50
6   TACATCTTCAGCCTGTTCGGCAAATGGAATAAATCTTAGTAGTCATCTGCCT 55
    |  ||  :   |||||:|| ||::|||   |||  |||::|| |||   |

50  uAspSerArgLeuHisAlaProMetTyrrPhePheLeuSerHisLeuAlaV 67
6   GGACCCCAAGCTGCACACCCCTATGTACTTCTTTCTTTCATCATCTGGGCCA 105
    |||   :||||||||||||:::|||||:|||:||||||||||||||||

67  alValAspIleAlaTyrrAlaCysAsnThrValProArgMetLeuValAsn 83
56  TTCTTGACATATCCTATAGCTTCCAACAAATGTCCCCAAGAAGTTGTCCAAC 105
    :::|||  ||||||  |  |:||||||:||||||:||||||||||||| ||

84  LeuLeuHisProAlaLysProIleSerProPheAlaGlyArgMetMetGlnrh 100
156 CTTATAAACCAGAAAAGAAACCCATCTCATTTGCCCATGCCATCACACACAC 205
    ||  ||| ||| || |  ||||||| |  ||||||||:|||||||:

100 rPheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValM 117
6   ATTCTTTCAGCCGCTTCGGAGTGCTTGAGTTTTGGCAGCCA 255
    |::|||||:|||||||:::|:|| |  |||  |||||||||:

117 etSerTyrAspLeuThrTyrVal.AlaIleCysHisProLeuArgTyrLeuAl 133
256 TGTCCTACGATACGGTTTGTGGCCCACTCTGCCACCCCTACACACACTGT 305
    ||||||||||   :|:  |||||||||||:||||||||||||||||| ::

133 aIleMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrT 150
306 CATCATGAGTTGGAAAGTATGTGTCGGCCCTGGCTGTCACATCTTGGTCAT 355

150 hrGlyValLeuLeuSerLeuLeuHisLeuValLeuLeuLeuProLeuPro 166
150 GTGGATTTAGCCTCTCTGTGGCACACAATTCTCCTGTAAGGTTGCCC 405
    |||||  |:||   |||:  |||||:|||||||:||||||||||||

167 PheCysArgProGlnLysIleThrTyrHisPhePheCysGluIleLeuAlaVa 183
406 TTTTGTGGGCCCCAGGAAAATAAACCACCCTCTTCTGTGAAAAATCCTGGCTGT 455
    ||||||||:|||||||| |||  ||||:|||:||||||||  |||||||

183 lLeuLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValValLeuA 200
456 TCTCAAGCTGGCCTGCGCGACACTTTGAATCAAACAAATGGTCATCCCTAG 505
    |||||||||||||||  :|:| |||||||:| |:|||:::|| | |

200 laGlyAlaIleSerGlyLeuValGlyProLeu.SerThrIleValAlaSe 216
506 CTGGCATGTCGTTTCGTCTTAGTGGACCCCCTTTGCCTTCAGTGCTTGTATC 555
    ||||  ||:|:|||:||||||||:|||||| ||||||||||||| ||||

216 rTyrMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValG 233
556 TTACACGGACATCCTCTGGACAAATCCTGAACAAATGCAGTCAAAAGAGGGTC 605
    ||| |  ||||||| ||||||||  ||||  ||||:::|||||||||||| ||

233 lnArgLysSerAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeu 249
606 GCAGAGAAGGCCTTTTCCACCTGTTCCTCCCATCCTGTGGGGGCTC 655
    ||||||   || ||||||||||||||   ||  ||| |||||||:

250 PheThrTyrGlyThrAlaIleIle.MetTyrVal.GlyProArgTyrGlyAsn 265
656 TTCTTTGCTATACCATGGCTTGGGCTTTATAGGTCCCGTGACTCTGAATCAG 705
    ||| |||  ||:||: |||::||:||  ||::|||||| |  |||:|| :||

266 ProLysGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPr 282
706 CGAGAGGAGGCAGGAGAGAAAATTTTGTCACTGTTCCACAGTCTTTTTACCCA 755
    ::|||||||:||||||||||||:::||:||| ||  ||||||||||  :||
```

```
282 oMetLeuAsnProLeuIleCysSerLeuArgAsn 293
                |||||||||||
756 AAGTGGAFTCCCTCATTAT...AGTCTGAGGAAC 786
```

seq_name: gb_est2:BF974217

seq_documentation_block:
```
LOCUS       BF974217       671 bp    mRNA    linear    EST 22-JAN-2001
DEFINITION  602243851F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4334965 5',
            mRNA sequence.
ACCESSION   BF974217
VERSION     BF974217.1  GI:12341432
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 671)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1205  row: e  column: 14
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1..671
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4334965"
                     /clone_lib="NIH_MGC_48"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      184 a    146 c    219 g    122 t
ORIGIN
```

alignment_block:
US-09-898-586-24 x BF974217/rev .

Align seq 1/1 to reverse of: BF974217 from: 1 to: 671

alignment_scores:
```
       Quality:     687.00      Length:      226
         Ratio:       3.734      Gaps:          7
Percent Similarity:   81.416    Percent Identity:   65.487
```

```
58 MetTyrPhePheLeuSerHisLeuAlaValAlaValAlaAspIleAlaTyrAlaCy 74
   ||||||||||       |||||||||||:|| |||||||
671 ATGTACTTCTTCTCTCACACCCTGGCGGTCGTCAACATCGCCTATGCCTG 622

74 sAsnThrValProArgMetLeuValAsnLeuLeuHisProAlaLysProI 91
   |||||||||||||||:::||||||||||||||||||| :|||||
621 CAACACACTGCCCCAGAGATGCTGGTGAACCTCCTGCATCCACCCAAGCCA 572

91 leSerPheAlaGly...ArgMetGlnThrPheLeuPheSerThrPhe 106
   |||||||||||||      :::      ||||||
```

571 TCTCCTTGCTGGCATGGCATGACAGTAGGAGGAGACCTTTCTTGCTGACTTTTT 522
```
107 AlaValThrGluCysSerLeuLeuLeuValValMetSerTyrAspLeuTyrVa 123
    |||||||||||:::    ||||:||||  ||||||||||||||:||
521 GCACATACTGAAATGCCCTCCTGTTGGTCGTCGATGGTCTTACGANCGGTACGT 472

123 IAlaIleCysHisProLeuArgTyrLeuLeuAlaIleMetThrArgValCys 140
    ||||||||:|||||||||:::||||||||||||||||||||||||||||
471 GGGCACTCTGCCACCCCTTCCGAPATTTCATCACGACCTGGAAAGTCT 422

140 ysIleThrLeuAlaValThrSerTrpThrThrGlyValLeuLeuSerLeu 156
    ||||||||||||||||||:::||||||||||||||||||||
421 GCATCACCTCTGGCCATCACTTCCTGGACACATGTGGCTCCTCCTGGCTATG 372

157 IleHisLeuValLeuLeuLeuProLeuProPheCysSArgProGlnLysSerIl 173
    :::|| |||||:::    |||||||||    :::|||||||
371 GTCCCATGTGAGCCGTCATCCTAAGACNCGCCCCTTTTGTGGGGCCTCGTGAAAT 322

173 eTyrHisPhePheCysSGluIleLeuLeuAlaValAlaValLeuLeu.LeuAlaCysAl 189
    |||||:::    ||||| |||||||||||||||||||||||||:::
321 CAACCACTTCTTCTGTGAAAATCCTGTCGTGTCTGCCTCAGGCCTGTGC 272

189 aAspThrHisIleAsnGlu...AsnMetValLeuAlaGlyAlaIleSerg 205
    ||||||||      :::     |||||||||||||||||||
271 TGATACCATGGCTCAACCAGGTAGGTCATCTTTGCAGCCGTGCCATGTCAT 222

205 lyLeuVal.GlyProLeuSerThrIleValValSerSerTyrMetCysIleIle 221
    |||||||| ||||:::    |||||||||||||||||||||:::
271 ACCTGGTAGGGACCACTCTGCCCGTGCCTGGTCTCCTACTCACACATCCT 172

221 u.CysAlaIleLeuGlnIleLeuGlnSerArgGluValGlnArgGlyLysAlaPhe 237
    :::||||||||||||||||||||||||||||      ||| |||||
221 ACTTGGGTAGGACCACTCTGCCCGTGCCTGGTCTCCTACTCACACATCCT 172

238 CysThr.CysPheSerHisSerLeuCysValIleGlyLeuPheThrTyrGlyThrA 254
    ||| ||||||||||||||||||||||||||||||||:::    |||||
171 AGGCGGCCATCCTGAGGAPCCAGTCTGGGGAGGGCCCGGAGAAAGGGCTTC 122

254 IaIleIleMetTyrValGlyProArgTyrGlyAsnProLysGluGlnLys 270
    ||||||||||||||||||||||||      :::     |||||
121 TCCACCATGGCTCCTCCACCTGCGTPAGTGGGCACTCTTCTTTGGCAGCCG 72

271 LysTyrLeuLeuLeuPhe 276
    ||| |||:::|||||
21 AAGGTCCTTTTTCTATTT 4
```

seq_name: gb_est2:BF116115

seq_documentation_block:
```
LOCUS       BF116115       732 bp    mRNA    linear    EST 24-OCT-2000
DEFINITION  7n77g04.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3570631 3',
            similar to TR:O95047 O95047 WUGSC:H_DJ0988G15.2 PROTEIN. ;, mRNA
            sequence.
ACCESSION   BF116115
VERSION     BF116115.1  GI:10985591
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 732)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D.
            R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
            Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
```

found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco
High quality sequence stop: 472.

FEATURES   Location/Qualifiers
source   1..732
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3570631"
                     /clone_lib="NCI_CGAP_Ov18"
                     /tissue_type="fibrotheoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGACGGGCGCGGCCGCACATTTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   152 a   217 c   145 g   218 t
ORIGIN

Align seq 1/1 to: BF116115 from: 1 to: 732

alignment_block:
US-09-898-586-24 x BF116115

alignment_scores:
      Quality: 678.00
      Ratio: 4.431
Percent Similarity: 88.953   Percent Identity: 79.070

  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
    ||||||::|| ||| ::||||||||||||||||||
219 ATGGGGGAAAATCAGACAAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 268

 17 oVal.GlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 33
    ::: ||||||||:|||||||:||||||:||::||||||||||||:||||
269 CCTGGGGCCCAAGGATTCAGATGCTCCTCCTTGGGCTCTTCTCCCTGTTC 318

 34 TyrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerIe 50
    ||:||||||:||||||||::||||||||||||||||||||||||||||||
319 TATATCTTCACCCTCCTGGGGAACGGGCCATCCTGGGGCTCATCTCACT 368

 50 uAspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaV 67
    |||||:||::|||||||||||::|||||||||||::|||||||:||||||
369 GGACTCCAGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCTG 418

 67 alValAspIleAlaTyrValAlaCysAsnThrValProArgMetLeuVaIAsn 83
    ||||:||||||||||||::::|||||||||||:|||||||||||::||||
419 TCGTCGACATCGCCTACACACCGCAACACGGTGCCCCAGATGCTGGCGAAC 468

 84 LeuLeuHisProAlaLysProIleSerPheLeuAlaGlyArgMetMetGlnThr 100
    ||||||||||||||||||||||||:||||:::|||||||||||||||||||:|
469 CTCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGCTCGCATGACGACAC 518

100 rPheLeuPheSerThrPheAlaVaIHisGluCysSerLeuLeuLeuValValM 117
    |||||||| ::|||:|||||||||||:|||||||||||||||||||:||
519 CTTTCTCGTTCTGTTTGAGTTTGGACACAGCGAATGTCTCCTGCTGGTGTGA 568

117 etSerTyrTyrAspleuTyrValAlaIleCysHisProLeuArgTyrLeuAla 133
    |||||||| |||:|||||||||||||||||:||
569 TGTCCTACGATCGTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTC 618

134 IleMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrTh 150
    ||||||||||||||||||||||||||||||||||||||||||||| |||
619 ATCATGACCTGGAGAGTCTGCATCACCCTGGCCGTTCACTCCTGACGTG 668

150 rGlyValLeuLeuSerIleLeuHisIleuValLeuLeuProLeuPro 167
    |||:||||:|:||||||||:::|| |||||:||||||||||||
150 | | | | | |||||| 167

alignment_scores:
      Quality: 643.00
      Ratio: 5.270
Percent Similarity: 100.000   Percent Identity: 100.000

Align seq 1/1 to reverse of: AI340119 from: 1 to: 367

alignment_block:
US-09-898-586-24 x AI340119/rev

BASE COUNT   115 a   73 c   116 g   63 t
ORIGIN

FEATURES   Location/Qualifiers
source   1..367
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1873767"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector. Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo."

Seq primer: -40UP from Gibco
High quality sequence stop: 245.

COMMENT   Unpublished (1997)
                     Contact: Robert Strausberg, Ph.D.
                     Email: cgapbs-r@mail.nih.gov
                     Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
                     Emmert-Buck, M.D., Ph.D.
                     cDNA Library Preparation: M. Bento Soares, Ph.D.
                     cDNA Library Arrayed by: Greg Lennon, Ph.D.
                     DNA Sequencing by: Washington University Genome Sequencing Center
                     Clone distribution: NCI-CGAP clone distribution information can be
                     found through the I.M.A.G.E. Consortium/LLNL at:
                     www-bio.llnl.gov/bbrp/image/image.html

JOURNAL   www-bio.llnl.gov/bbrp/image/image.html
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                     Tumor Gene Index
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
REFERENCE   1 (bases 1 to 367)
                     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM   Homo sapiens
SOURCE   human.
KEYWORDS   EST.
VERSION   AI340119.1   GI:4077046
ACCESSION   AI340119
                     ;, mRNA sequence.
                     similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
DEFINITION   qk64f08.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1873767 3'
LOCUS   AI340119   367 bp   mRNA   linear EST 29-DEC-1998

seq_documentation_block:
seq_name: gb_est1:AI340119

669 TGGCTCCTCCTCCTGGCTCCTGGCGCCCCATGGTGGTTCTCATCCTAAGACTGCCCT 718
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
167 heCysArgProGln 171
    |||||
719 TCTCTGGGCGCCTCAA 732

 60 PhePheLeuSerHisLeuAlaValAlaValAspIleAlaTyrValAlaCysAsnTh 76
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
367 TTCTTCCTCTCACACCCTGGCCGTCGACAACCGCTACGCCTACGCCTGCAACAC 318

 76 rValProArgMetLeuValAsnLeuLeuHisProAlaLysProIleSerP 93
    |||||||||||||||||||||||||||||||||||||||||||||||||||

seq_name: gb_est1:AI285267

```
317  GGTGCCCCGGATGCGTGGTGAACCTCGATCCAGCCAGCCATCTCCT  268
93   heAlaGlyArgMetMetGlnThrPheLeuPheSerThrPheAlaValThr  109
     |||||||||||||||||||||||||||||||||||||||||||||||||
267  TTGCGGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACA  218
110  GlyCysLeuLeuLeuLeuValMetSerTyrAspLeuTyrValAlaIleCy  126
     ||||||||||||||||||||||||||||||||||||||||||||||
217  GAAGTGTCTCCTCTGGGTGGTGAGTCCTATGATCTGTACGTGGCCATCTG  168
126  sHisProLeuArgTyrLeuAlaIleMetThrThrProArgValCysIleThr  143
     |||||||||||||||||||||||||||||||||||||||||||||||||
167  CCACCCCTCCGATATTTGGCCATCATGACCACTGGAGTCTGCATCACC  118
143  euAlaValThrSerTrpThrThrGlyValLeuLeuSerLeuHisLeu  159
     ||||||||||||||||||||||||||||||||||||||||||||
117  TCCGGGTGACTTCCTGACCACTGGAGTCCTTTTATCCTTGATTCATCTT  68
160  ValLeuLeuLeuProLeuProPheCysArgProGlnLysIleTyrHisPh  176
     |||||||||||||||||||||||||||||||||||||||||||||||
67   GTGTTACTTCTACCTTTACCTTCTGTAGGCCCCAGAAAAATTTATCACTT  18
176  ePheCysGluIleLeu  181
     ||||||||||||||||
17   TTTTTGTGAAATCTTA  2
```

seq_documentation_block:

```
LOCUS        AI285267          361 bp    mRNA      linear    EST 29-JAN-1999
DEFINITION   ql38c08.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1874606 3',
             similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
             ; , mRNA sequence.
ACCESSION    AI285267
VERSION      AI285267.1  GI:3923500
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 361)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
             Insert Length: 659     Std Error: 0.00
             Seq primer: -40UP from Gibco
             High quality sequence stop: 26.
FEATURES             Location/Qualifiers
     source          1..361
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1874606"
                     /clone_lib="NCI_CGAP_Co8"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     colon adenocarcinoma, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector. Library is normalized. Library was constructed by
```

BASE COUNT   115 a   73 c   110 g   63 t
ORIGIN   Bento Soares and M. Fatima Bonaldo. "

alignment_scores:
```
                Quality: 602.00        Length:  120
                Ratio:   5.190         Gaps:    0
Percent Similarity: 96.667     Percent Identity: 95.833
```

alignment_block:
US-09-898-586-24 x AI285267/rev ..

Align seq 1/1 to reverse of: AI285267 from: 1 to: 361

```
60   PhePheLeuSerHisLeuAlaValValAlaValAspIleAlaValTyrAlaCysAsnTh  76
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
361  TTTTCCCTCTCACACCTGGCGTGGTGAAGTCCAATGAATTTGTTGTTGACATGGCCTACAAC  312
76   rValProArgMetLeuValAlaSnLeuLeuHisLeuProAlaLysProIleSerP  93
     ||||||||||||||||||||||||||||||||||||||||||||||||||
311  GGTGCCCCGGATGCGTGGTGAACCTCCTGCATCTGCCAATGGCCATCACCC  262
93   sHisProLeuArgTyrLeuAlaIleMetThrThrProArgValCysIleThr  143
     ||||||||||||||||||||||||||||||||||||||||||||||||
143  euAlaValThrSerTrpThrThrGlyValLeuLeuSerLeuHisLeu  159
     ||||||||||||||||||||||||||||||||||||||||||||
261  TTGCGGGGCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGCTGTCACA  212
110  GlyCysLeuLeuLeuLeuValMetSerTyrAspLeuTyrValAlaIleCy  126
     ||||||||||||||||||||||||||||||||||||||||||||||
211  GAAGTGTCTCCTCTGGGTGGTGAGTCCTATGATCTGTACGTGGCCATCTG  162
126  sHisProLeuArgTyrLeuAlaIleMetThrThrProArgValCysIleThr  143
     ||||||||||||||||||||||||||||||||||||||||||||||||
161  CCACCCCTCCGATATTTGGCCATCATGACCACTGGAGTCTGCATCACC  112
143  euAlaValThrSerTrpThrThrGlyValLeuLeuSerLeuHisLeu  159
     ||||||||||||||||||||||||||||||||||||||||||||
111  TCCGGGTGACTTCCTGACCACTGGAGTCCTTTTATCCTTGATTCATCTT  62
160  ValLeuLeuLeuProLeuProPheCysArgProGlnLysIleTyrHisPh  176
     |||||||||||||||||||||||||||||||||||||||||||||||
61   GTGTTACTTCTACCTTTACCTTCTGTAGGCCCCAGAAAAATTTATCACTT  12
176  ePheCysGlu  179
     ||||||||||
11   TTTTTGTGAA  2
```

seq_name: gb_est2:H62441

```
LOCUS        H62441           426 bp    mRNA      linear    EST 06-OCT-1995
DEFINITION   yr79d08.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
             IMAGE:211103 5' similar to SP:OLF7_MOUSE P34984 OLFACTORY
             RECEPTOR-LIKE PROTEIN K7 ; , mRNA sequence.
ACCESSION    H62441
VERSION      H62441.1  GI:1015273
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 426)
AUTHORS      Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
             Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
             ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
             ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
             Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
             Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE        Generation and analysis of 280,000 human expressed sequence tags
JOURNAL      Genome Res. 6 (9), 807-828 (1996)
MEDLINE      97044478
COMMENT      Contact: Wilson RK
             Washington University School of Medicine
```

```
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1867
High quality sequence stops: 304
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1867      Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 304.

FEATURES             Location/Qualifiers
Source               1. .426
                     /organism="Homo sapiens"
                     /db_xref="GDB:3784344"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:211503"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT        75 a    140 c    95 g    111 t    5 others
ORIGIN
```

```
alignment_scores:
        Quality:      593.00          Length:       131
        Ratio:        4.669           Gaps:         2
Percent Similarity:   96.947      Percent Identity: 96.947

alignment.block:
US-09-898-586-24 x H62441

Align seg 1/1 to: H62441    from: 1  to: 426
```

```
                AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProValGl  19
  3  AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProValGl
     |||||||||||||||||||||||||||||||||||||||||||||||||||
  1  GACAATATAACAATCCATCACAGAGTTCCTCCTACTGGGAATTTCCGGTTGG  50

     yProArgIleGlnMetLeuPheGlyLeuPheSerLeuPheTyrValP  36
 19  yProArgIleGlnMetLeuPheGlyLeuPheSerLeuPheTyrValP
     |||||||||||||||||||||||||||||||||||||||||||||||
 51  CCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTTCCCGTGTTCTACGTCT  100

     heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer  52
 36  heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer
     ||||||||||||||||||||||||||||||||||||||||||||||||||
101  TCACCCTCTGGGACACGGGACCAATACTGGGGCTCATCTCACTGGACTCC  150

     ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeu.AlaValValA  69
 53  ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeu.AlaValValA
     ||||||||||||||||||||||||||||||||||||||||  |||||||||
151  AGACTGCACGCCCCCATGTACTTCTTCCTCACACCTGGGCGGTCGTCG  200

     spIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeu  85
 69  spIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeu
     ||||||||||||||||||||||||||||||||||||||||||||||||||
201  ACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACCTCCTG  250

     HisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLe 102
 86  HisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLe
     ||||||||||||||||||||||||||||||||||||||||||||||||||
251  CATCCAGCCAAGCCAATCCTCTTTGCGGGCCG.ATGATGCAGACCTTTCT  299

     uPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerT 119
102  uPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerT
     ||||||||||||||||||||||||||||||||||||||||||||||||||
300  GTTTTCCACTTTTGCTGTCACAGAAGTGTCTCCTCCGGTGTGAGTCCT  349

     yrAspLeuTyrVal.AlaIleCysHisProLeuArgTyr 131
119  yrAspLeuTyrVal.AlaIleCysHisProLeuArgTyr
     ||||||||||||||  ||||||||||||||||||||||||
350  ATGATCTGTACGTNGGCCATCTGNCACCCCCTCCGATAT 388
```

OM of: US-09-898-586-24 to: Issued_Patents_NA:* out_format : pfs

Date: Aug 21, 2002   8:19 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV-xlh
-Q=/cgn2_1/USPTO_spool/US09898586/runat_16082002_2057735_20570/app_query.fasta_1.373
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH-0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP-4.500 -QGAPEXT-0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT-0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX-100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US0909898586_@CGN1_1_48 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXX -WAIT -THREADS=1

Search information block:
Query: US-09-898-586-24
Query length: 310
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 45.940000

score_list:

Sequence	Strd Orig ZScore	EScore	Len	Documentation		
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-948A-1	+	1048.50	2174.83	1713	1.9e-113	1713
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-467-947A-1	+	1048.50	2174.83	1713	1.9e-113	1713
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-5	+	609.00	1254.94	966	3.3e-62	966
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-827-291A-1	+	605.00	1245.13	1290	1.1e-61	1290
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-7	+	592.00	1219.09	966	3.2e-60	966
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-8	+	578.00	1189.56	966	1.4e-58	966
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-748-506-6	+	574.00	1181.13	966	4.2e-58	966
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-085-371-5	+	543.00	1116.61	900	1.7e-54	900
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-748-506-9	+	452.00	923.64	984	9.3e-44	984
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-07093-1	+	369.00	743.75	1474	9.7e-34	1474
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-439-313-526	+	384.00	780.51	963	8.7e-36	963
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-988-876-2	+	382.50	769.63	1838	3.5e-35	1838
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-465-980-1	+	369.00	743.75	1474	9.7e-34	1474
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-053-303-1	+	369.00	743.75	1474	9.7e-34	1474
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-07093-1	+	369.00	743.75	1474	9.7e-34	1474
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-706-281A-17	+	185.50	361.73	978	1.8e-12	978
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-06352-84	-	181.00	348.63	1320	9.9e-12	1320
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-097-231-17	+	185.50	361.73	978	1.8e-12	978
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-671-525B-9	+	182.50	355.44	975	4.1e-12	975
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-672-109B-9	+	182.50	355.44	975	4.1e-12	975
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-9	+	182.50	355.44	975	4.1e-12	975
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-238-9	+	182.50	355.44	975	4.1e-12	975
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-629-335B-9	+	181.50	352-84	975	4.1e-12	975
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-599-252-84	-	181.00	348.63	1320	9.9e-12	1320
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-436-074-57	-	181.00	348.63	1320	9.9e-12	1320
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US96-06352-84	-	181.00	348.63	1320	9.9e-12	1320
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-870-511-7	+	175.00	339.34	999	3.8e-11	999
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-672-109B-7	+	174.50	337.34	1080	4.2e-11	1080
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-7	+	174.50	337.34	1080	4.2e-11	1080
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-842-238-5	+	174.50	337.34	1080	4.2e-11	1080
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-5	+	174.50	337.34	1080	4.2e-11	1080
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-629-335B-5	+	174.50	337.34	1080	4.2e-11	1080
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-629-335B-5	+	174.50	337.34	1080	4.2e-11	1080
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-671-525B-7	+	173.00	335.15	996	5.6e-11	996
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-842-045-7	+	173.00	335.15	996	5.6e-11	996
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-842-238-7	+	173.00	335.15	996	5.6e-11	996
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-629-335B-7	+	173.00	335.15	996	5.6e-11	996
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-842-238-7	+	173.00	335.12	996	5.6e-11	996
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-870-511-9	+	173.00	335.12	999	1.2e-10	999
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-870-511-11	+	173.00	335.12	999	1.2e-10	999
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-662-560-1	+	173.00	328.92	1671	1.2e-10	1671
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-780-749A-5	+	173.00	328.92	1671	1.2e-10	1671

/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-870-511-5 + 173.00 328.92 1.2e-10 1671
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-686-591-3 + 169.00 316.60 6.0e-10 2308
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-976-715-3 + 169.00 316.60 6.0e-10 2308
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-706-281A-15 + 165.00 312.05 1.1e-09 1671
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-097-231-15 + 165.00 312.05 1.1e-09 1671

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-948A-1

seq_documentation_block:
; Sequence 1, Application US/08467948A
; Patent No. 5998164
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; TITLE OF INVENTION: Coupled Receptor GPR2
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,948A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116..1003
US-08-467-948A-1

alignment_scores:
Quality: 1048.50   Length: 305
Ratio: 4.017   Gaps: 4
Percent Similarity: 85.574   Percent Identity: 69.180

alignment_block:
US-09-898-586-24 x US-08-467-948A-1

Align seq 1/1 to: US-08-467-948A-1 from: 1 to: 1713

  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
    ||| ||| ||| ||| ||| ||| ||: ||||||||||||||||||||||

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-467-947A-1

seq_documentation_block:
; Sequence 1, Application US/08467947A
; Patent No. 6090575
GENERAL INFORMATION:
APPLICANT:
APPLICANT: LI, YI
APPLICANT: CAO, LIANG
APPLICANT: NI, JIAN
APPLICANT: GENTZ, REINER
APPLICANT: BULT, CAROL J.
APPLICANT: SUTTON III, GRANGER G.
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
TITLE OF INVENTION: Coupled Receptor GPR1
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,947A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04079
FILING DATE: 30-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 116..1003
US-08-467-947A-1

alignment_scores:
Quality: 1048.50   Length: 305
Ratio: 4.017   Gaps: 4
Percent Similarity: 85.574   Percent Identity: 69.180

alignment_block:
US-09-898-586-24 x US-08-467-947A-1

Align seq 1/1   to: US-08-467-947A-1   from: 1 to: 1713

          1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProPr 17
            ||| ||| ||| |||||||||||||||||::||||||||||||||||||
        128 ATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 177

         17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhePr 34
            ::||||||||||||||||||||||:::||||||||||||||||||||||||
        178 CCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTTCCCCTGTTTCT 227

Patent No. 6159707

GENERAL INFORMATION:
  APPLICANT: Ronnett et al.
  TITLE OF INVENTION: NOVEL SPERM RECEPTORS
  NUMBER OF SEQUENCES: 31
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Leydig, Voit & Mayer, Ltd.
    STREET: Two Prudential Plaza, Suite 4900
    CITY: Chicago
    STATE: IL
    COUNTRY: US
    ZIP: 60601-6780
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/748,506
    FILING DATE: 08-NOV-1996
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 60/033,751
    FILING DATE: 09-NOV-1995
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    REFERENCE/DOCKET NUMBER: 74940
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 312-616-5600
    TELEFAX: 312-616-5700
INFORMATION FOR SEQ ID NO: 5:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 966 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
US-08-748-506-5

alignment_scores:
  Quality: 609.00   Length: 308
  Ratio: 2.819   Gaps: 0
  Percent Similarity: 70.130   Percent Identity: 40.909

alignment_block:
US-09-898-586-24 x US-08-748-506-5

Align seq 1/1   to: US-08-748-506-5   from: 1   to: 966

3 AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProValGl 19
  :::||| :||||| ||:||
28 GAGAATAGTTGACTGTCAAACACTTTGCATTTGCCAAGTTCTCTGAGGT 77

19 yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheThrTyrValp 36
   ||| ||||||||||:||||::|:: |||:::||:::
78 CCCTGGAGAATGCTTCCTCCCGTTCAACCTCATCCTTCATGTTCTTAG 127

36 heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52
   ::||||||||||:||||||||:||||::||||: :::
128 TATCACTAACAGGAAATACTCCATAGTCCTTGCTATTTGTACCAGTCCA 177

53 ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValValAs 69
   |||||||:|||||||||||||||||||:|:::::||::
178 TCTCTACACACCCCCATGTACTTCTTTCTTCGCCCAACTTGTCTCCTGGA 227

69 pIleAlaTyrIleAlaCysAsnThrValProArgMetLeuValAsnLeuHis 86
   ::|||:|||||:||:||||||:::::||||||||:::::
228 GATTGGCTATACTTGCTTCTGTCTCATACCCAAGATGCTGCAGAGCCTTGTGA 277

86 isProAlaLysProIleSerPheAlaGlyArgMetGlnThrPheLeu 102
   ||||:::|||||||| |||:::
278 GTCAGGCCAGAGAGAATCTCGGGAGGATGTGCCACACAGATGTTTTTTT 327

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-748-506-5
; Sequence 5, Application US/08748506
seq_documentation_block:

34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyIleLeuSerLeu 50
   |||||||||||||||||||||||||||||||||||||||||
228 ATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGGGGATCCTCACTG 277

51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
   ||||||||:|||||||:::|||||||||||||||||||||||
278 GACTCCAGACTCCACACCCCCATGTACTTCTTCCTCTCACACCTGGCCGT 327

67 lValAspIleAlaTyrIleAlaCysAsnThrValProArgMetLeuValAsnL 84
   ||||:::|||||||:||:||||||||:::::||||||:||
328 CGTCAACAATCGGCTATGCCTGCAACAACAGTGCCCAGAATGCTGGTGAACC 377

84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr 100
   ||||:::||||||||:::|||||||| |||:::
378 TCCTGCATCCACCCAAGCCCATCTCCTTTGCTGGTCGCATGACACTAGAC 427

101 PheLeuPheSerThrPheAlaValThrGlyCysLeuLeuLeuValValMe 117
    :::|||||||:::|||||||||:||||||||||||||||||:|||:||
428 TTTCTCTTTTGAGTTTTGCACATACTGAAAGCCTCCTGTTGGTGGTCATG 477

117 tSerTyrTyrAspLeuTyrValAlaIleCysHisLeuProLeuArgTyrLeuAlaI 134
    ||||||| :::||||||||:||:|||||||||||||:::|||:::| |
478 GTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTCATCA 527

134 leMetThrThrArgValCysIleThrThrLeuAlaValAlaValThrTrpThrThr 150
    ||||||||||||||||||||||||||||||||||||||||
528 TCATGACCACTCGAAAGTCTGCATCACTACTCTGGCCGTAGCAGTGACATGT 577

151 GlyValAlaIleSerLeuIleHisSerLeuValLeuLeuLeuProLeuProHis 167
    |||||||||||||::||||||||||||||||||||||||||
578 GGCTCCCTCTCCCTGGCTAYGGCCATGGAGCCCATCCTAAGACTGCCCTT 627

167 eCysArgProGlnLysSerIleThrTyrHisPhePheCysGlnIleLeuAlaVal 184
    ||||||||||:|||:::||||||||||||||:|||||||:::||
628 TTGTGGGCCTTCCGTGAAAATCAACCACTTCTTCTGTGAAATCCTGGCTGTCC 677

184 euLysLeuLeuAlaCysAlaAspThrHisIleLeuAsnGluAsnMetValLeuAla 200
    :::||:||||||||||||||||||:::||||||||||||||:::
678 TCAGGCTGCTGGCCTGTGCTGATACCCACAGTGGTCATCTTTGAA 727

201 GlyAlaIleLeuSerGlyLeuValGlyProLeuSerPheThrIleValValSerTy 217
    ||:|||||||||||||||||||::||||::||||||||||||||||
728 GGCTGCAATGTTCATCCTGGTTGGGACCACTGTCTGGGTGTCTCCCTA 777

217 rMetCysSerIleLeuCysSerAlaIleIleLeuGlnIleSerArgGluValGln 234
    ||||| |||||:||||:::||||||||||||:::
778 CTCACACATTCCTCTGGGGGGGCAATCCAGTCTGGGGGAGGGGCCCA 827

234 rgLysAlaLeuSerThrCysSerPheSerHisLeuLeuValGlyLeuPhe 250
    |||||||||||||||||||||||||||||||||::||||
828 GAAAGGCGCTTCCCACCTGCGTCCCCACCCTCGGGTAGTGGGACTCTTC 877

251 TyrGlyThrAlaIleIleLeuMetThrTyrValGlyValAsnProLy 267
    :|:||||||||||||||||||||||||||||||:::||:::
878 TTTGGSAGCGCCCATCGTTCAGTACAYAYGGCCCCTAAGTCCGCCATCTGA 927

267 sGlyGlnIleLysTyrLeu.LeuLeuPheHisSerLeuPheAsnPro... 282
    :::|||||:||||| ||: |||:||||||||||
928 GGAGCAGCAGAAGGTCCTTTTTCTTATTTTACAGTTCCTTTCAACCCGA 977

283 MetLeuAsnProLeuIle.CysSerLeuArgAsnSerGlu.ValLysAsn 298
    ||||||||:::::||||| ||::|||||::: ||| |||||::
978 TGCTTAAACCCCCTGATTTACAACCCTGAGGAATGTAGAGGGTCAAGGGT 1027

299 ThrLeuLys 301
    :::\|\|:::
1028 GCCCTCCGA 1036

```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-827-291A-1

seq_documentation_block:
; Sequence 1, Application US/08827291A
; Patent No. 5874243
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Sathe, Ganesh
;   TITLE OF INVENTION: NOVEL OLRCD15 RECEPTOR
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: SmithKline Beecham Corporation
;     STREET: 709 Swedeland Road
;     CITY: King of Prussia
;     STATE: PA
;     COUNTRY:
;     ZIP: 19406
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
```

```
103 PheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerTy   119
    |||::| |||||||||||||||||||||||||||||||||||||||||||
328 TTCGCATTTTTTGTGATAACTGAGTCGTGCCTATTGGCCATGGCCTT      377

119 rAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMetT   136
    :|||   ||||||||||||||||||||||||||||||||||||| |||
378 TGACCGCCTGCATGGCCATATGCCTCCCCACTCCACTATGCAACCCGAATGA  427

136 hrTrpArgValCysIleThrLeuAlaValThrThrSerTrpThrThrGlyVal 152
    ||||||||||||||||||||||||||||||||||||||||||||||||
428 GTCGNGAGGTATGTCCCCATTTGCCAATTGTTTTCATGGGAAATGGGAATGC  477

153 LeuLeuSerLeuLeuHisLeuValLeuLeuLeuProLeuProPheCysAr   169
    ::::||||||||||||||||||||||||||||||||||||||||
478 ATAGTAAGCTCGGACAAACCAAATTTATTTTCTCCTTGAACTTCGTGG      527

169 gProGlnLysIleTyrHisPhePheCysGlnIleLeuAlaValLeuLyst   186
    ||||      ||||||||||||||||||||||||||||||
528 ACCCTGAAATAGACCACTTCTTCTGTCACCTTCCACCTCTCCTGGCAC      577

186 euAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAla   202
    |||||||||||||||| |||||
578 TTGCCCTGTGGAGAATACATCCCAAAACGAGGCTGCCATCTTTGTGGTAGCA 627

203 IleSerGlyLeuValGlyProLeuSerThrIleValSerTyrMetCy      219
    ||||||||||||||||||||||||||||||||||||||||
628 GTCCTCTGCAATAATCTAGCCCATTTTTGCTGATCATTTATTCTTATGTCAA 677

219 sIleLeuCysAlaIleLeuGlnIleGlnSerArgGlnValGlnArgLysA   236
    |||  |||||||||||||||||:||||
678 AATTCCTCATTGCAGTGCTNCTGATGCCTTCACCTGAGGGGCCCCATAAG   727

236 laPheCysThrCysPheSerHisLeuCysValIleGlyIleLeuPheTyrGly 252
    ||||||  ||||||||||||||||||||||||||||||||||
728 CTCTTTCCACCTGTTCGTCTCACCTACTTGTAGTCACACTTTTTTATGGC   777

253 ThrAlaIleIleMetThrTyrValGlyProArgTyrGlyAsnProLysGluGl 269
    |||| |||| |||||||||||||||||||||||||||||||||
778 TCAGCATGTATTACCTATTTGAGCCCCAAGTCTAGCCCACTCACCAGGAAT   827

269 nLysLysTyrTyrLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP   286
    :||||||||||||||||||||||||||||||||||||||
828 GGACAAAATTCTTGGCCCTCTTCTACACAGTAGTGACATCCACATTGCAACC 877

286 roLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArg   302
    |||||||||||||||||||||||||||||||||||||||
878 CTATCACTCTATAGTTTAAGGAACAAGGAAGTCAAGGACAGCACTGAGAAGA 927

303 ValLeuGlyValGlyLeuArgAlaLeu                          310
    |||||||||||||||||||||::
928 ACTCCTGGGCCTGAAAAAATTCTG 951
```

```
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/827,291A
;   FILING DATE: 28-MAR-1997
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: King, William T
;   REGISTRATION NUMBER: 30,954
;   REFERENCE/DOCKET NUMBER: GP50001
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 610-270-5015
;   TELEFAX: 610-270-5090
;   TELEX:
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1290 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
; US-08-827-291A-1

alignment_block:
US-09-898-586-24:
US-08-827-291A-1

Align seq 1/1   to: US-08-827-291A-1   from: 1   to: 1290

alignment_scores:
       Quality:          606.00           Length:     304
       Ratio:            2.742             Gaps:       2
Percent Similarity:  72.697   Percent Identity: 40.132
```

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-748-506-7

alignment_scores:
```
                        Quality:     592.00            Length:    308
                        Ratio:        2.779            Gaps:        0
Percent Similarity:     69.156     Percent Identity:  40.260
```

alignment_block:
US-09-898-586-24 x US-08-748-506-7 ..

Align seq 1/1 to: US-08-748-506-7 from: 1 to: 966

```
152 ValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCy 168
    ::::|::::     |||    :::  ::|:         :::|
752 TCTACAGATGGAATCATTTAYGCTGTAGCCACATTTTCCTTCCTTACTG 801

  3 AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePheLeu 19
    ::::|::    ::||||  :|::::||||||         |||
 28 GAGAATAGTTTGTCTGTGTCAAAGCGTTTGCAATTTGCCAAGTTCTCYAGGT 77

168 sArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValLeuL 185
    |      ||||||||:|||||||         |||||         :|
802 TGGCTCTCGGGAAATAGCCCACTTCTTCTGTGAGTTACCTTCCCTACTAA 851

 19 yProArgIleGlnMetLeuuPheGlyLeuPheSerLeuPheTyrValP 36
    |||         ::||||||||||         |||:|:|::     |
 78 CCCTGGAGAAATGCTTCCTCCTCCGYTCACCCCTCATCCTTCTCAGTTCTTAG 127

185 ysLeuLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGly 201
    ||:::|||         |:||||         :||:::::::
852 TCCTTCTCAYGCAAATGACACATCAAATATTTGRAAAGGTYATTTCATYTGGC 901

 36 heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52
    ||:||         :::||||||:::||         |||||:|:|:::
128 TACAACTAACAGGAAAATACTCTCATAGCCCTTGCTATTTGTTGACCAGTCCA 177

202 AlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTyrMe 218
    |||||         |||||||:::||  :::|||||:||:||         
902 TCTATAAGTAAATGGCTTTGTTTTTCCCTGTTGCAATCATCAYTTGCTTCCTTATGC 951

 53 ArgLeuHisIleAlaProMetTyrPhePheLeuSerHisLeuAlaValAs 69
    |:||||:||  ::|||||||||::||||:||         :::|:|:::
178 TCTCTACACAGCCCCCATGTACTTCYTTCTGGCCAACTGTCYCTCCCTGGA 227

218 tCysIleLeuCysAlaIleLeuIleGlnIleSerArgGlyLValGlnArgL 235
    ||:::|||         |:||||  :::||||||:::|||||  :::|||
952 TGGAGTTATTCTGGCTGYCATTCACATGGAANTCTGGAGAGGGTCGTCGCA 1001

 69 plIeAlaThrYrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuH 86
    |:|||||:::|::||||  ||:|:||  :::|||||::||||||:|||
228 GAATGGCYATAYACTTGCYCYGYCATAYCCCAAGATGCTGCAGAGCCYTGYGA 277

235 ysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPheTyr 251
    ||||:|||||:||||::||:||||||         :::|:||:|||||
1002 AAGCTTTCACGACCTGTTCCTCTCACCTCATGGTGGTGGGGAATGTTCTAT 1051

 86 isProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeu 102
    |:||||:|::|  :::||||||         ||||||||:||||||
278 GTGAGGCCCGAGGGAGCATCTCTTGGGAGGGTTGYGCCTCACAGATGYCTTC 327

252 GlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLysGl 268
    |||||||         |:||||||         :||||||||
1052 GGACAGGGTTGTTCATGTACAYCACGCCACATTCTGATCCGCTCCCCAAC 1101

103 PheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerTy 119
    ::|||  :|||:||||  :::||||||||||  ::::||:|||:::
328 TTCATATTCTTTGGTAYAACTGAGTGCTGCCTYATYCGGCAGCCATGGCCYT 377

268 uGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetLeuA 285
    |||         |:|||         |:||||||||:::||||
1102 GCAGGACAAAGCTGGTGTCGYTGYATTCTACACCATCCTCACTCCCATGCTGA 1151

119 rAspLeuTyrValAlaIleGlyCysHisProLeuuArgTyrLeuAlaIleLeuMetT 136
    |||         |||||||||:||:::||||||:::|||||||||::
378 TGACCGCYATAYGCCTAYAYGTYCCCCACTCCACTAYGCAAACCGAAATGA 427

285 snProLeuIleCysSerLeuArgAsnSerGlyValLysAsnThrLeuLys 301
    :::||||||         :::||  :|||||||||  |::||||||||
1152 ATCCCCTCAYTCTACAGCCTCCGCCAACAAAGGAAGYGACCAGAGCAYTCAYG 1201

136 hrIleuProArgValCysIleuThrLeuAlaValIleThrSerTrpThrThrGlyVal 152
    :::|||||||||||         |||:|:||||||:||
428 GTCGTGGGGYGTAYGYGCYCAYTTGCAATTGYCTCAYGGGGTYGATGGGATGC 477

302 ArgValLeuGly 305
    |||     |||
1202 AAGATCTCAGGA 1213

153 LeuLeuSerLeuIleLeuHisLeuValLeuLeuLeuProLeuProPheCysAr 169
    |||||:|||         ||:::||||||||||::||||
478 ATACTAGGYCYTCYGGGACAGACCAAATTTATYTCYCCYTGAACTTCYGTGG 527

169 gProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValLeuLeuLysL 186
    |||         ||:||||||||||         :::|||||||||||||:
528 ACCCCTGYGAGAYAGAGCCACACTTCTTCTGYGAYCTTCCACCCTCCYGGCAC 577

186 euAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAla 202
    ||||||||         :::||:::||:::||||||:|:|||||||
578 TTGCCTGYGTGTGATACAYCCCAAAATCAGGCCATCATYTTGYGTGGCAGCA 627

203 IleSerGlyLeuValGlyProLeuSerThrIleValValSerTyrMetCy 219
    :::|||         |::||||||||  :|::|||:||:|::|
628 GTCCTCTCCATATTTAGTCCATYTYTYACTGAYCATTTCYTYCCTATGTCAG 677
```

seq_documentation_block:
    Sequence 8, Application US/08748506
;   Patent No. 6159707
; GENERAL INFORMATION:
;   APPLICANT: Ronnett et al.
;   TITLE OF INVENTION: NOVEL SPERM RECEPTORS
;   NUMBER OF SEQUENCES: 31
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Leydig, Voit & Mayer, Ltd.
;   STREET: Two Prudential Plaza, Suite 4900
;   CITY: Chicago
;   STATE: IL
;   COUNTRY: US
;   ZIP: 60601-6780
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/748,506
;   FILING DATE: 08-NOV-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/033,751
;   FILING DATE: 09-NOV-1995
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   REFERENCE/DOCKET NUMBER: 74940
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312-616-5600
;   TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 966 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-748-506-8

alignment_scores:
    Quality: 578.00   Length: 308
    Ratio: 2.714   Gaps: 0
    Percent Similarity: 69.156   Percent Identity: 38.961

seq_name: /cgn2_6/ptcodata/2/ina/6A_COMB.seq:US-08-748-506-8

```
219 sIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnArgLysA 236
    |||||| |||  |:|||
678 AATTCCGTTCCAGTGCCTGGTGATGCCTTCACCTGAGGGGCCCATAAAG 727

236 laPheCysThrCysPheSerHisLeuCysValIleGlyTyrLeuPheTyrGly 252
    |||:| ||||| ||||
728 CTCTCCCTACCTGTTCATCTCACCTACGTTAGTCACACTCTTCTATGGC 777

253 ThrAlaIleIleMetThrTyrValGlyProArgTyrGlyAsnProLysGluGl 269
    |||:|| ||||||
778 TCAACATCTGCACCTATTTGAGGTCCAAGTCTAGCCACTCACCAGGAGT 827

269 nLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP 286
    ||| ||| ||
828 GGACAAACTCCTTGGCCCTTCTATACATCACTGACATCCCATGCTGAATC 877

286 roLeuIleCysSerLeuArgAsnSerGluValIleValLysAsnThrLeuLysArg 302
    |||:|| |||||||||||||:||||
878 CCATCATCACAGCTTAAGGAACAAGAAGTAAAAGGTGCACTGAGAAGA 927

303 ValLeuGlyValValGlnArgAlaLeu 310
    :||||||
928 ACTCTGGCCCTGAACAAAGTTCTG 951
```

alignment_block:
US-09-898-586-24 x US-08-748-506-8

Align seg 1/1 to: US-08-748-506-8 from: 1 to: 966

```
  3 AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProValGl 19
    ||| :|: ||
 28 GAGAATAGTTGTGTCGTCAAAACGTTTTGCATTTGCCAAGTTCTCTGAGGT 77

 19 yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrValP 36
    :||| ||||| |
 78 CCCTGGAGAAGTGCTTCCTCCTGTTCACCCTGTTCAGCCTGTTCTATGT 127

 36 heThrLeuLeuGlyAsnGlyThrIleLeuPheLeuGlyLeuIleLeuSerLeuAspSer 52
    |||| | :||| ||
128 TATCACTAACAGGAAAATGGCTCTCAATGCCCTTGCTATTTTGTTACCAGTCCA 177

 53 ArgLeuHisAlaProMetThrTyrPhePheLeuSerHisLeuAlaValAlaS 69
    |||| |||::|| |||||
178 TCTCTACACACCCCCATGTACTTCTTTCTGGCCAACTTGTCTCCTGGA 227

 69 pIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuH 86
    :||:: ||| ||
228 GATTGGCATAACTTGGCCTGTAATAACCCAAGATGCTGCAGAGTCTTGTGA 277

 86 isProAlaLysProIleSerProPheAlaGlyArgMetMetGlnThrPheLeu 102
    ||:: :|| |||
278 GTGAGGCCCGAGACAGATCTTTCAGGTGGGATGTGCCACACAGATGTTTTTC 327

103 PheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerTyr 119
    |||| |||| |||
328 TTCATATTCTTTGGTGTATAACTGAGTGTGCCTTATTGGCACCATGGCCTT 377

119 rAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMett 136
    ||| |:|||||
378 TGACCCCTATATGGCTATATTGTTCCCCACTCCACTARGCAACCCGAATGA 427

136 hrTrpArgValCysSerIleThrLeuAlaValThrSerTrpThrThrGlyVal 152
    |||| |||:||
428 GTCCGTGAGGTATGTGCCCACTGGCAAATTGTTCATGGGTGATGGGATGC 477

153 LeuLeuSerLeuIleHisLeuValValLeuLeuProLeuProPheCysAr 169
    ||| || :||
478 ATAGTAGGTCTGGGACAGACCAATTTATTTCTCCTTGAACTTCTGTGG 527

169 gProGlnLysIleTyrHisPhePheCysGluIleLeuLeuAlaValLeuSL 186
    ||| ||||||
528 ACCCTGTGAGATAGACCACTTCTTCTGTGATCTTCACCCTCCGGCAC 577

186 euAlaCysAlaAspThrHisSerIleAsnGluAsnMetValLeuAlaGlyAla 202
    |||:::||||
578 TTGCCTGTGGTGATACATCCAAAATGCAGGCTGCCAATCTTTGTGGTAGTT 627

203 IleLeuSerGlyLeuValGlyProLeuSerHrIleValValSerTyrMetCy 219
    ||||:| ||
628 GTCCTCTGCATATCTAGCCCCTTTTTGCTGATCATTTATTGTTATGTCAG 677

219 sIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnArgLysA 236
    |||||| |||:||
678 AATTCCGTTCCAGTGCCTGGTGATGCCTTCACCTGAGGGGCCCATAAAG 727

236 laPheCysThrCysPheSerHisLeuCysValIleGlyTyrLeuPheTyrGly 252
    |||:||||||| ||||
728 CCCTTTCAAACCTGTTCCTCCACCTACTGTAGTCACACTCTTTTATGGC 777

253 ThrAlaIleIleMetThrTyrValGlyProArgTyrGlyAsnProLysGluGl 269
    ||| :|||||||
778 TCAGGATCTGTTACCTATTTGAGGTCCAAGTCTAGCCACTCACCAGGAAT 827

269 nLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP 286
    ||| |||||||||
828 GGACAAACTCCTTGGCCCTTCTACACAGCACAGTGACATCCATGTTGAACC 877
```

Left column:

```
286 roLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArg 302
    |:::||:|    ||||||||||||||||||:||:::||||||
878 CTATCAACTCATAGTTTAAGGAACAAGGATGTCAAGGCACGACACTGAGAGA 927

303 ValLeuGlyValGluArgAlaLeu 310
878 ATTCTGGCCCTGAAAAAAATTCTG 951
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-6

; seq_documentation_block:
; Sequence 6, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

alignment_scores:
    Quality: 574.00
    Ratio: 2.720
Percent Similarity: 68.506   Percent Identity: 39.286

alignment_block:
US-09-898-586-24 x US-08-748-506-6   .

Align seg 1/1 to: US-08-748-506-6   from: 1 to: 966

```
  3 AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProValG1 19
    ::|||   ||| ||||||||||||:||||
 28 GAGAATAGTTGTCTGTCAAACGCTTTGCATTTGCCAAGTTCTCTGAGGT 77

 19 yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrValP 36
    ||||
 78 CCCTGGAGAATGCCTTCCTCCTGTTCACCCTCATCCTTCCTTCTCATTGT 127

 36 heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52
    :::|||    |||||||:::|||
128 TATCACTAACAGGAAAATGCTCTCATAGCCCTTGCTGNTTGTACCAGTCCA 177
```

Right column:

```
 53 ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValValAs 69
    ||||||||::::||||||||||||||||||::::||||:||:::||||
178 TCTCTACACACCCCAMGTACTTCTTTCTTGGCCAACCTTGTCTCCTGGA 227

 69 pIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuH 86
    ||||    |||:||||||||||:::||||||||||||:||:::|||
228 GATTGGCTATACTGGCTCGTCATACCAAGATACCAGAGATGCTGCAGAGCCTTGTGA 277

 86 isProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeu 102
    ||||   ||||||||    |||||||||||||||||||:::||:::
278 GTGAGGGCCAGAGAGATCTCAGGGAGGGATGTGCCACACAGATGTTTTTC 327

103 PheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerTy 119
    |||||   |||||||||    |||||||||:::|||||||
328 TTCACATTTTTTGGCATAAACTGAGTGCTGCCTATTGGCACCATGGCCTT 377

119 rAspLeuThrValAlaIleIleCysHisProLeuArgTyrLeuAlaIleMetTr 136
    ||||   |||||:|||||||||:::||||||||||||||:::
378 TGACCCGTGCAATGGGCATATAGTCTCCCCACTTCCCACTATGCAACCCGAAATGA 427

136 hrrpArgValCysIleThrLeuAlaValThrSerTrpThrThrGlyVal 152
    ||||    |||||||||||||||||:::|||||||||
428 GTCGTGAGGTAGTGGCCATTTGGCAAATTGTTTCATMGGAGAAATGGGATGC 477

153 LeuLeuSerLeuIleHisLeuValLeuLeuLeuLeuProLeuProPheCysAr 169
    ::|||::|||||:::|||    |||||||||:::||||||
478 ATAGTAGGTCTGGGACAGAGCCAAATTNAATTNCTTCCCTTGAACTTCTGTG 527

169 gProGlyLysIleTyrHisPhePheCysGluIleLeuAlaValLeuLysL 186
    ||||    |||:||||||||||||||::::||||||    |
528 ACCTTGTGGAGATAGACCCACTTCTTCTGTGACCTTCCACCTCTTGGCAC 577

186 euAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAla 202
    |||||||||:::|||   ||||||:::||||||
578 TTGCCCTGTGGTGATACATCCCAAAACGAGGGTGCCATCTTTGGTGGCAGCA 627

203 IleSerGlyLeuValGlyProLeuSerThrIleValValSerTyrMetCy 219
    ||||   :::|||||||||||:::|||||||
628 ATCCCTCGTGTATATCTAGTCCATTTTTGGGTGATCCTTTTATTGTTAGTCAG 677

219 sIleLeuCysAlaIleIleLeuGlnIleSerArgGluValGlnArgLysA 236
    ||||||||    |||||||||||||||||:::|||||
678 AATTCTCGTTGCAGTGCTGGTTGGATGCCTTCACCTGAGGGCGGCCANAAAG 727

236 laPheCysPheSerSerIleLeuCysValIleGlyLeuPheTyrGly 252
    |::|||    |||||||:::|||||||
728 CTCTTTCCACCTGTTCCTCACACCTACTTGTTAGTCACACTCTTTTTATGGC 777

253 ThrAlaIleIleMetTyrValGlyProArgThrIleGlyAsnProLysGluGl 269
    :::|||   ||||||:::|||||||
778 TCTGTGGTCCTTTACCTATTTGAGGGCCCAAGTCTAGCCACTCACCAGGAAT 827

269 nLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP 286
    :::|||    ||||||:::||||||||:::
828 GGACAAAACTCTTGGCCCTCTTCTACACAGCAGTGACATCCATGCTAACC 877

286 roLeuIleCysSerLeuArgAsnSerGluValValLysAsnThrLeuLysArg 302
    |:::||:|    ||||||||||||||||||:||:::||||||
878 CTATCAACTCATAGTTTAAGGAACAAGGATGTCAAGGCACGACACTGAGAGA 927

303 ValLeuGlyValGluArgAlaLeu 310
    :::|||   :::|||:::
928 ACTCCGACCTGAAAAAAATTATG 951
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-085-371-5

; seq_documentation_block:
; Sequence 5, Application US/09085371
; Patent No. 6218358
; GENERAL INFORMATION:
; APPLICANT: Firestein, Stuart
; APPLICANT: Zhao, Haiqing

' TITLE OF INVENTION: Functional Expression of, and Assay for, Functional Cellular Re
' TITLE OF INVENTION: vivo
' FILE REFERENCE: P0151US2 / 09805059
' CURRENT APPLICATION NUMBER: US/09/085,371
' CURRENT FILING DATE: 1998-05-19
' PRIOR APPLICATION NUMBER: US 08/891,243
' PRIOR FILING DATE: 1997-07-10
' PRIOR APPLICATION NUMBER: US 60/045,961
' PRIOR FILING DATE: 1997-05-07
' NUMBER OF SEQ ID NOS: 6
' SOFTWARE: Patentin version 3.0

' SEQ ID NO 5
' LENGTH: 900
' TYPE: DNA
' ORGANISM: Rat
US-09-085-371-5

alignment_scores:
        Quality: 543.00
Percent Similarity: Ratio: 2.675
                                 70.000   Percent Identity: 39.310

alignment_block:
US-09-898-586-24 x US-09-085-371-5

Align seq 1/1   to: US-09-085-371-5   from: 1   to: 900

                                                                    Length: 290
                                                                    Gaps: 1

8   IleThrGluPheLeuLeuLeuGlyPheProValGlyProArgIleGlnMe 24
    :::|||||||:|||::||||||||||||::::||||||||:::||||:
28 GTGAGTGAATTTGTGTTGGTTGGCTGGGTTTCCCCAGGTCCCCACTGCGAGT 77

24 tLeuLeuPheGlyLeuPheSerLeuPheTyrValPheThrIleuLeuGlyA 41
    :::||||||::|||||||::||||||||||:::||||||||||
78 ACTACTATTTTCCTTTCTCTCTGGACTATGTGTTGGTTGGTTGACTGAAA 127

41 snGlyThrIleLeuGlyIleSerLeuAspSerArgLeuHisAlaPro 57
    :::||||:|| ||| ::||| ||||||||
128 ACATCCTCATCATTATTAGGAACCACCCAACCTCCACAAACCC 177

58 MetTyrPhePheLeuSerHisLeuAlaValAlaValAlaAspIleAlaTyrAlaCy 74
    ||||||||||||:||::|||:::||||||||::|||| ||:
178 ATGTATTTTTTTTCTTCGCCTAATATATGTCATTTCTCGAGATTTGGTATGTCAC 227

74 sAsnThrValProArgMetLeuvValAsnLeuLeu........His 87
    |||:::|||:::|||:|||| :::|||:||: |||
228 TGTTACGGATTCCTAAGATCCTGCGCTTCAATTGGTTCCAAGGAGAAACC 277

87 roAlaLysProIleSerPheAlaGlyArgMetGlnThrPheLeuPhe 103
    ::::: ||||/|||| |||/ || ::::|||||
278 ATGGACAGCTGAATCTCCTTTGAGGCATGCCATGACAACTCTACTTTTTC 327

104 SerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerTyrAs 120
    :::::: |||||||||||||::::||||||||||::||||||
328 CTGGGCTTGGGTTGGCACAGAGTGTCCTTCTTGCGTGATGGCTATGA 377

137 rpArgValCysIleThrLeuAlaValThrSerTrpIleThrGlyValLeu 153
    :::::||:|||||:::|||||||||||||||:||: :::|||
378 CCGGCTATGTGGGCTATCTGCTACCACCAACCCCGTCAATTGTCAGTA 427

120 pLeuThrValAlaIleCysHisProLeuAArgTyrLeuAlaIleMetThrT 137
    |||||||:|:::||||||:::||||||||::||| :::||
428 GCCGGCTATGTGTGCCAGAAGCGCAGCTGGATCCTGGGACGGTTTTGGT 477

154 LeuSerLeuIleHisLeuValLeuProLeuProPheCysArgPr 170
    :::||||:::|||||||||||||||||:::||| :::|||
478 ATCTCCATGGTAAAGTTTTCCTTATTTCTCGCCTGTCTTACTGTGGCCC 527

170 oGlnLysIleTyrHisPhePheCysGluIleLeuAlaValLeuLysLeuA 187
    ::::: |||||||||||||::::|||||||||||::||||
528 CAACACCATCAACCACTTTTCTGTGAGTGTCTCCAATTGCTCAACCTGT 577

187 laCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAlaAlaIle 203

---

578 ::|||:::||| :::|| |||
    CATGCACTGACATGTCCACAGCAGAGCTTACAGACTTTGTCCTGGGCCATT 627

578 SerGlyLeuValGlyProLeuSerThrIleValValSerTyrMetCysIl 594
    :::|||||||||||::||||||||||:::||||||||||||::
    TTTATTCTGTGCCTGGGACCGGCTCTCTCTCACTGGGGGCATCCTACATGGGGCAT 677

220 eLeuCysAlaIleLeuGlnIleGluSerArgGluValGlnArgLysAlaP 237
    ||:::|||||||:::|||||/ ||||| ::::|||||||
    CACAGGTGCTGTGATCGGATCCCCTCACCTGCTGGGGCGGCGGCAAAAGCCT 727

237 heCysThrCysPheSerHisSerLeuCysSerValIleGlyLeuPheTyrGlyThr 253
    |:::||||||| :::|||||||||||| :::|||||||::
    TTTCAACCTGTGCCTCCCACCTCACTGTTGTGATCATCTTCTTCTTAATGCAGCC 777

254 AlaIleLeuMetTyrValGlyProArgTyrGlyAsnProMetLeuAsnProL 270
    :::::|||||||||:::|||||||::::||||||||::::|||
    AGCTGGTCTCTGTACTCTACGCTGTCAATGTACCGGTTGTTCAATCCCA 827

270 SLysThrValLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnProL 287
    :||| ||:::|||::::|||||::|||||||||||::::
    CAAGCTGGTCTCTGTACTCTACGCTGTCAATGTACCGGTTGTTCAATCCCA 877

287 euIleCysSerLeuArgAsn 293
    :::||||||||||||||
    TCATCTACTGCTTGCGCAAC 897

---

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-748-506-9

seq_documentation_block:
' Sequence 9, Application US/08748506
Patent No. 6159707
' GENERAL INFORMATION:
' APPLICANT:
' APPLICANT: Ronnett et al.
' TITLE OF INVENTION: NOVEL SPERM RECEPTORS
' NUMBER OF SEQUENCES: 31
' CORRESPONDENCE ADDRESS:
' ADDRESSEE: Leydig, Voit & Mayer, Ltd.
' STREET: Two Prudential Plaza, Suite 4900
' CITY: Chicago
' STATE: IL
' COUNTRY: US
' ZIP: 60601-6780
' COMPUTER READABLE FORM:
' MEDIUM TYPE: Floppy disk
' COMPUTER: IBM PC compatible
' OPERATING SYSTEM: PC-DOS/MS-DOS
' SOFTWARE: Patentin Release #1.0, Version #1.25
' CURRENT APPLICATION DATA:
' APPLICATION NUMBER: US/08/748,506
' FILING DATE: 08-NOV-1996
' CLASSIFICATION: 435
' PRIOR APPLICATION DATA:
' APPLICATION NUMBER: US 60/033,751
' FILING DATE: 09-NOV-1995
' CLASSIFICATION: 435
' ATTORNEY/AGENT INFORMATION:
' REFERENCE/DOCKET NUMBER: 74940
' TELECOMMUNICATION INFORMATION:
' TELEPHONE: 312-616-5600
' TELEFAX: 312-616-5700
' INFORMATION FOR SEQ ID NO: 9:
' SEQUENCE CHARACTERISTICS:
' LENGTH: 984 base pairs
' TYPE: nucleic acid
' STRANDEDNESS: double
' TOPOLOGY: linear
' MOLECULE TYPE: DNA (genomic)
US-08-748-506-9

alignment_scores:
        Quality: 452.00   Length: 301

Left column

Percent Similarity: 67.774   Percent Identity: 32.226

Ratio: 2.216   Gaps: 0

alignment_block:
US-09-898-586-24 x US-08-748-506-9
US-08-748-506-9   from: 1 to: 984

Align seg 1/1 to: US-08-748-506-9 from: 1 to: 984

```
  4  AsnIleThrSerIleThrGluPheLeuLeuGlyPheProValGlyPr   20
     |||:  :||  |||::||::||| :::|||| |||||::  ||:  |::
 31  AATGGGACTTTGGTCCTGGAGTTCAATCCTTGAGGGGTACCCTGGGCCGA  80

 20  oArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrValPhe   37
     ||:  | ||:||  | |:::| |||   |||  ||:  ||:  |:|
 81  GCACCTGAAGATCCTTCTTCTTCCTACTGACTTGCTTGGCTGGGGCCT  130

 37  hrLeuLeuGlyLysAsnGlyThrIleLeuGlyLeuIleLeuSerLeuAspSerArg  53
     ::|||::|||:|||||  || ||| ::|| ||| |:::|||||||||::
131  CCCTCATGGGCAACATGCTCATAATTACCACTGTGTGGACCACCGA  180

 54  LeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValValAspIl   70
     ::|  :||| ||||||::|||||:|||||:|||||:::||||||:||
181  CTTGCAGACGCCCATGTACTTCTTTCTCACCACCTTCTCTTTTGTGGAGTG  230

 70  eAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuHisP   87
     ||| |||| |||||:||| |||||  |:||||| |:||||  |:|||
231  TTGTTTTATAACTACTGCTAATCCCCAGCTCCTCACCATCATTTGTCAG  280

 87  roAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeuPhe  103
     |||:  :| ||||| |||:||  :::|| |||||| :::::|| |
281  GGAGGCAAAAGATTCCCTTTGGGGTCGTTACGTGTCTTATTTTGGGCTTA  330

104  SerThrPheAlaValThrGluCysLeuLeuValValMetSerTyrAs    120
     ::|  ||| |||||| ||||||::|||||:|||||::|||||:
331  CTTGTCGTGGGGGCAACAGGTTTTTCCTTTGGGTGCCGTTATCCGTGGA  380

120  oGlnLysIleTyrHisProLeuArgTyrLeuAlaIleMetThrT       137
     :::||||  |::|::||| ||| | |||:| |::|||||||:|
381  CCGCTTCTTCGGCCATCTGCAAACCTTCTACATTATCCAACCATCATTGAGCC  430

137  rpArgValCysIleThrLeuAlaValThrSerTrpThrThrGlyValLeu  153
     | |||:|||| |:|| |||:::|||:::||| :||| |||:||:
431  CAAGGATGTGGCTTCCTTCCTGGTTACGTGTCGTTATTGTTTTGGGCTTCC  480

154  LeuSerLeuLeuHisLeuValLeuLeuProLeuProCysArgPr        170
     ::|  :::||||||||::|| |||::||||||::||| |::|
481  TTCATGGCCAGTCCAGTTGTCATGGTCTTTCCAAGACATTTTACTGTGGTCC  530

170  oGlnLysIleTyrHisPheCysGluIleLeuLeuAlaValLeuLysLeuA  187
     |||::||||||:||| ||| ||:::||::||||:|||| ::::|:
531  AAACATTATTCCTCACTTTTTCTGATTTGGACCACTGGCAAATCTCT  580

187  laCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAlaIle  203
     ||||| ||||:||||||  |:::::|||| |:||| |::|::
581  CCTGTTCAGAAAACCAGGTCTATTGAGATGCTGTTTTTTTTTTTTTACCCTTGCTGTA  630

204  SerGlyLeuValGlyProLeuSerThrIleValSerArgTyrMetCysIl  220
     ||| | :::||| |:| :|| |:|||  :|||||| |  ||
631  ATTGTCCTTTTTGCTCCTTTTGCATACACAGCAAATAT  680

220  eLeuCysAlaIleLeuLeuGlnIleLeuSerArgGluValGlnArgLysAlap  237
     ||:|  ||||||||::|:|| | |:|| ||| ||| ||:|||||
681  AGTAGTCACCATAGTCAGAGACTCCCTTCCAGGGAGCGACAGAGAGCTT  730

237  heCysThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThr  253
     |  |||::|:|| |||||::|||| |||::|| |:|:|||||:||||:
731  TTCCACCTGCTCCTCTCCCACCTCAATGTGCTGCTCTCTCTAATGTATGGCAGC  780

254  AlaIleIleMetThrValGlyProTyrGlyAsnProLysGluGlnLy     270
     ::|:|::|::|||| ||||||::|||||| |||||:|||||:
781  TGTGCATTTATATACCTGAAGCCAAAGCAGAAGCACAGAGACTGGACCAA  830

270  sLysTyrLeuLeuPheHisSerLeuPheAsnProMetLeuAsnProL    287
     | ::||:|| ||:|||||:||||||| |||| | |||||||  |
781  TGTGCATTTATATACCTGAAGCCAAAGCAGAAGCACAGAGACTGGACCAA  830
```

Right column

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-439-313-526

seq_documentation_block:
Sequence 526, Application US/09439313
Patent No. 6329505

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqui
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 526
LENGTH: 963
TYPE: DNA
ORGANISM: Homo sapiens
US-09-439-313-526

alignment_scores:
Quality: 384.00   Length: 320
Ratio: 1.939   Gaps: 8
Percent Similarity: 61.875 Percent Identity: 31.250

alignment_block:
US-09-898-586-24 x US-09-439-313-526

Align seg 1/1 to: US-09-439-313-526 from: 1 to: 963

```
  4  AsnIleThrSerIleThrGluPheLeuLeuGlyPheGlyPro........   17
     ||:|:  ||                                  .||||
 13  AACTTCACACACATGGCCACC...TTTGTGCTTATTGGTATCCCAGGATTAGA  59

 18  ............ValGlyProArgIleGlnMetLeuLeuPheGlyL     29
                 |||||                          |||  |
 60  GAAAGCCCCATTTCTGGGTGGC.....................TTCCCC   88

 29  euPheSerLeuPheTyrValPheThrLeuLeuGlyLeuAsnGlyThrIleLeu  45
     ||:||||  ||:::||  |||||||  |||||||||||||||
 89  TCCTTTCCATG.....TATGTAGTGGCAATGTTTGGAAACTCCATCCGTGGTC  135

 46  GlyLeuIleLeuSerLeuAspSerArgLeuHisAlaProMetTyrPhePheLe  62
     ||  ||||||||||||||::||| ||:||||||  |||  |::|||:|
136  TTCATCGTGAAGGACGGAACGCAGCCTGCACGCCTCCGATGTACCTCTTTCT  185

 62  uSerHisLeuAlaValValAspIleAlaTyrAlaCysAsnThrValProA  79
     :||::||||:|||||||||||::|||  ||| ||||||||  |:||:
186  CTCCACGCCTTGCAGCCATTGACCTGCCTTATCCACACACCAATCACCATGTCCTA  235

 79  rgMetLeuValAsnLeuLeuHisProAlaLysProIleSerPheAlaGly  95
     |::||::|||||:::|||||::|| |||  ||||:::
186  CTCCACGCCTTGCAGCCATTGACCTGCCTTATCCACACACCAATCACCATGTCCTA  235
```

```
831  CAGAGAGGCTGCTGCTCTGTTGTGAACATGGTTGTGACACCCTTCTGAACCCTG  880

287  euIleCysSerLeuArgAsnSerGluValValLysAsnThrLeuLysArgVal  303
     ::|||  ||||||||||||||:::||:|:||| |||||||||||||:
881  TCATCTACACACCCTGGCCAACAAAGCACAGGTCCACCAGGCTCTCAGGGATCCT  930

304  Leu 304
     :||
931  CTG 933
```

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-988-876-2

seq_documentation_block:
Sequence 2, Application US/08988876
Patent No. 6063596
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
TITLE OF INVENTION: WITH IMMUNE RESPONSE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,876
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0441 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1828 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNOT01
CLONE: 364702
US-08-988-876-2

alignment_scores:
Quality: 382.50   Length: 307
Ratio: 1.922   Gaps: 8
Percent Similarity: 64.821   Percent Identity: 37.785

alignment_block:
US-09-898-586-24 x US-08-988-876-2

Align seg 1/1 to: US-08-988-876-2 from: 1 to: 1828

```
seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-465-980-1

seq_documentation_block:
  Sequence 1, Application US/08465980
  Patent No. 5756309
  GENERAL INFORMATION:
;   APPLICANT: Soppet, Daniel R.
;   APPLICANT: Li, Yi
;   APPLICANT: Rosen, Craig A.
;   APPLICANT: Ruben, Steven M.
;   TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;     ADDRESSEE: STUART & OLSTEIN
;     STREET: 6 Becker Farm Road
;     CITY: Roseland
;     STATE: New Jersey
;     COUNTRY: USA
;     ZIP: 07068
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
```

```
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/465,980
;     FILING DATE: 06-JUN-1995
;     CLASSIFICATION: 536
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Ferraro, Gregory D.
;     REGISTRATION NUMBER: 36,134
;     REFERENCE/DOCKET NUMBER: 325800-446
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 201-994-1700
;     TELEFAX: 201-994-1744
;   INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1474 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;       NAME/KEY: CDS
;       LOCATION: 274..1233
US-08-465-980-1

alignment_block:
US-09-898-586-24 x US-08-465-980-1 .

alignment_scores:
  Quality:  369.00       Length:      320
  Ratio:      1.883      Gaps:          8
  Percent Similarity: 61.250     Percent Identity: 30.625

Align seq 1/1   to: US-08-465-980-1    from: 1   to: 1474
```

(Sequence alignment data follows)

```
seq_name: /cgn2_6/ptodata/2/ina/5b_COMB.seq:US-09-053-303-1

146 ThrSerTrpThrThrGlyValLeuLeuSerLeuIleHisLeuValLeuIle 162
        ||||: .: |::
709 GTGGCTGTGGTCCGCGGAGTCCCGTCTTTTTTTTTCCCACTGCCTCTGTGAT 758

162 uLeuProLeuProPheCysArgProGlnLysIleTyrHisPheCysG 179
        ||||| :|||
759 CAAGCGGGCTGGCCTTCTCGCCACTCCAARGTCCTCTCGCCACTCCTATTGTG 808

179 luIleLeuAlaValLeuLysLeuAlaCysAlaAspThrHisIleAsnGlu 195
        |||:|::|| |:
809 TCCACCAGGATGTAATGAAGTTGGCCTATGCAGACACT............. 846

196 AsnMetValLeuAlaGlyAlaIleSerGlyLeuValGlyProLeuSerTh 212
        |::|||::::     |::|||
847 .........TTGCCCAATGTGGTATANTGGTCTTACTGCCATTCTGCTGGT 887

212 r........................IleValValSerTyrMetCysIleLeuLeuC 222
        ||| |:|||||||||:.:
888 CATGGGCGTGGACGTAATGTCATCTCCGTTGCTATTTTCTGATAATAC 937

222 ysAlaIleLeuGlnIleGlnSerArgLeuValGlnLnArgLysAlaPheCys 238
        ::|||:|||||||| | ||||||| |||||||||
938 GAACGGTTCTGCAAACTGCCTCCAAGTCAGAGGCGCCAAGGCCTTTGGA 987

239 ThrCysPheSerHisLeuCysSerValIleGlyLeuPheTyrGlyThrAlaI 255
        ||||| :|:: ||||||| ||| :
988 ACCTGTGTGTCACACATTGGTGTGGTCACTCGCGCCTTCTATGGTGACACTAT 1037

255 eIleMetTyrValGlyProArgTyrGlyAsnProLysGluGlnLysLysT 272
        :.:| ||| ||| |:
1038 TGGCCTCTCAGTTGTACACGGCGTTTGGAAACAGGCCTTCATCCACATTGTGC 1087

272 yrLeuLeuLeu.......PheHisSerLeuPheAsnProMetLeuAsnPro 286
        :.:::::::       |:|||||||| ::|||:||||
1088 GTGTTGTCATGGGTGACATCTACCTGCTGCCCTGTGTCATCAATCCC 1137

287 LeuIleCysSerLeuArgAsnSerGlyValLysAsnThrLeuLysArgVa 303
        ||:| |||| | ||||:||:||| ||:||
1138 ATCATCTATGGTGCCAAAAACACAGATCAGAACA.........CGGGT 1178

303 lLeuGlyVal 306
        |||| |
1179 GCTCGCTATG 1188
```

```
seq_documentation_block:
Sequence 1, Application US/09053303
Patent No. 5948890
GENERAL INFORMATION:
    APPLICANT: Soppet, Daniel R.
    APPLICANT: Li, Yi
    APPLICANT: Rosen, Craig A.
    APPLICANT: Ruben, Steven M.
    TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
    NUMBER OF SEQUENCES: 8
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
        ADDRESSEE: STUART & OLSTEIN
        STREET: 6 Becker Farm Road
        CITY: Roseland
        STATE: New Jersey
        COUNTRY: USA
        ZIP: 07068
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/053,303
        FILING DATE:
        CLASSIFICATION:
```

```
    PRIOR APPLICATION DATA:
        APPLICATION NUMBER: US 08/465,980
        FILING DATE: 06-JUN-1995
    ATTORNEY/AGENT INFORMATION:
        NAME: Ferraro, Gregory D.
        REGISTRATION NUMBER: 36,134
        REFERENCE/DOCKET NUMBER: 325800-446
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: 201-994-1700
        TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 1474 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    FEATURE:
        NAME/KEY: CDS
        LOCATION: 274..1233
US-09-053-303-1

alignment_scores:
        Quality:    369.00                    Length:  320
        Ratio:        1.883                    Gaps:      8
    Percent Similarity: 61.250     Percent Identity: 30.625

alignment_block:
US-09-898-586-24 x US-09-053-303-1

Align seq 1/1  to: US-09-053-303-1  from: 1  to: 1474 .

  4 AsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePro........... 17
        ||||:||:||| | :::||::|: |:.::  |:
286 AACTTCACACATGCCACC...TGTGTGCTTATTGGTATCCCAGGATTAGA 332

 18 ..........ValGlyProArgIleGlnMetLeuLeuPheGlyL 29
             ||||| |
...                                       TTCCCCC 361

 29 euPheSerLeuPheThrTyrValPheThrArgLeuLeuGlyAsnGlyThrIleLeu 45
362 TCCTTTCCATG...TATGTPAGTPGGCAAATGTGTGGAAAACTGCATCGTGGTC 408

 46 GlyLeuIleLeuSerLeuAspSerArgLeuHisAlaAproMetTyrPhePheF 62
409 TTCATCGTAAGGACGGAACCGCAGCCTGCAGCGCTCCGATTGAGTACCTCTTTCT 458

 62 uSerHisLeuAlaValAlaValAspIleAlaTyrAlaCysAsnThrValProA 79
459 CTGCANGCTTCCAGCCATTGAACCTGGCCTTATCCACATCCACATGCCTA 508

 79 rgMetLeuValAsnLeuLeuHisProAlaLysProIleSerPheAlaGly 95
509 AGATCCTTGCCCTTTCTGGTTTGAATTCCGAGAGATTAGCATTGAGGCC 558

 96 ArgMetMetGlnThrPheLeuPheSerThrPheAlaValThrThrGluCysIle 112
559 TGTCTTPACCCAGAGAGTTCTTTAATTCATGCCCTCTCAGCCATTGAATCCAC 608

112 uLeuLeuValValMetSerThrTyrAspLeuThrTyrValAlaIleCysHisProL 129
609 CATCCTGCTTGGCCATGGCCTTTGACCATTGTCAGCATCAGGCCATCTGCCACCAC 658

129 euArgTyrLeuAlaIleMetThrThrTpArgValCysIleThrLeuAlaVal 145
659 TCCGCCATGGCTGCAGTGCTCAACAATACAGTAACAGCCCAGATTGGCATC 708

146 ThrSerTrpThrThrGlyValLeuLeuSerLeuIleHisLeuValLeuIle 162
709 GTGGCTGTGGTCCGCGGAGTCCCGTCTTTTTTTTTCCCACTGCCTCTGTGAT 758
```

seq_name: /cgn2_6/ptcdata/2/ina/PCTUS_COMB.seq:PCT-US95-07093-1

seq_documentation_block:
; Sequence 1, Application PC/TUS9507093
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.

REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-446
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 201-994-1700
  TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1474 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  FEATURE:
  NAME/KEY: CDS
  LOCATION: 274..1233
PCT-US95-07093-1

alignment_block:
US-09-898-586-24 x PCT-US95-07093-1

Align seq 1/1   to: PCT-US95-07093-1   from: 1   to: 1474

alignment_scores:
  Quality: 369.00   Length: 320
  Ratio: 1.883   Gaps: 8
  Percent Similarity: 61.250   Percent Identity: 30.625

4   AsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePro......... 17
    |||:||||:||:||| :::||||||||||||||||
286 AACTTCACACACATGCCACC...TGTGTGCTTATTGGTTA
CCAGGATTAGA 332

18 ....ValGlyProArgIleGlnMetLeuLeuPheGlyI   29
                         :::||||||||||||||||
333 GAAAGCCCATTTCTGGGTTGGC.....................TTCCCCC 361

29 euPheSerLeuPheTyrValPheThrLeuLeuGlyAsnGlyThrIleLeu   45
    ||:|||| ||||| ::||||:|||||||::||
362 TCCTTTCCATG...TATGTAGGTGGCAAATGTGGAAACTGCATCGTGGTC 408

46 GlyLeuIleLeuSerLeuAspSerArgLeuHisAlaProMetTyrPheIle   62
    |||:|||||||||||| |||| |||||||||:||||||
409 TTCATCGTAAGGACGGAACCCAGCCTGCACGCTCCGATGTACCTCTTTCT 458

62 uSerHisLeuAlaValAlaAspIleAlaTyrArgAlaCysAsnThrValPro   79
    ||:|||||| ||||:||||| :::||||||||:|||||||
459 CTGCATGCTTGCAGCCATTGACCTGGCCTTATCCACATCCACATGCCCTA 508

79 rgMetLeuValAsnLeuLeuHisProAlaLysProGlIeSerPheAlaGly   95
    ||:||||| |||:|||:|||| |||||||
509 AGATCCTTGCCCTTTTCTGGGTTGGATTCCCGGAGAGATTAGCAATGAGGCC 558

96 ArgMetMetGlnThrPheLeuPheSerThrPheAlaValThrGluCysSer   112
    ||||| ||||| ||| |||||||:||
559 TGTCTTACCCAGAGATGTTCTTTATTCATGCCCCTCCAGCCATTGAAATCCAC 608

112 uLeuLeuValValMetSerTyrArgAspLeuTyrValAlaIleCysHisProL   129
    ||||||| |||:|||| :||||||||||||
609 CAVCCTCCTGGGCATGGCCATGCCGTAATGTGGCCAATCTGCCACGGCAC 658

129 euArgTyrLeuAlaIleMetThrTrpArgValCysIleThrLeuAlaVal   145
    ||||||||||| ||| :|||||||:::||||||||
659 TGCCCCATGCTGCAGTGCTCAACAATACAGTAACAGCCCAGATTGGCATC 708

146 ThrSerTrpThrThrGlyValLeuLeuuSerLeuIleHisLeuValValLeu   162
    ||||||| |||:|||| |||||:||||||
709 GTGGCCTGTGGCGTCCGGGGATCCCTCTTTTTTTTTCCCACTGCCCTGTGAT 758

162 uLeuProLeuProPheCysArgProGlnLysIleTyrHisPhePheCysG   179
    ||| ||||||| ||||||| :||||||||
759 CAAGCGGCGTGGCCTTCTGCCAACTCCAAATGTCCTCCGCACTCCTATTGTG 808

179 luIleLeuAlaValLeuLysLeuAlaCysSerAlaAspThrHisIleAsnGlu 195

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-706-281A-17

seq_documentation_block:
Sequence 17, Application US/08706281A
Patent No. 6100048
GENERAL INFORMATION:
  APPLICANT: Cone, Roger D
  APPLICANT: Fan, Wei
  APPLICANT: Boston, Bruce A
  APPLICANT: Kesterton, Robert A
  APPLICANT: Lu, Dongsi
  APPLICANT: Chen, Wenbiao
  TITLE OF INVENTION: Methods and Reagents for Discovering and
  TITLE OF INVENTION: Using Mammalian Melanocortin Receptor Agonists and Antagonists
  TITLE OF INVENTION: To Modulate Feeding Behavior in Animals
  NUMBER OF SEQUENCES: 19
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
    STREET: 300 South Wacker Drive
    CITY: Chicago
    STATE: IL
    COUNTRY: USA
    ZIP: 60606
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/706,281A
    FILING DATE: 04-SEP-1996
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: No. 61000048nan, Kevin E
    REGISTRATION NUMBER: 35,303
    REFERENCE/DOCKET NUMBER: 96,886
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 312-913-0001
    TELEFAX: 312-913-0002
    TELEX:
INFORMATION FOR SEQ ID NO: 17:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 978 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 1..975
  US-08-706-281A-17

alignment_scores:
  Quality: 185.50   Length: 305
  Ratio: 1.124   Gaps: 13
  Percent Similarity: 54.098   Percent Identity: 24.918

alignment_block:
US-09-898-586-24 x US-08-706-281A-17

Align seq 1/1 to: US-08-706-281A-17   from: 1 to: 978

```
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
    ||||||||||||::: |:||ss:::::||:|
706 GGGCTATTACC...CTCACCATGCTACTGGGGATTTCATTGTCTGCTG 752

217 rMet...CysIleLeuCysAlaIleLeuGlnIleSerArgGluValG 233
    ::|||
753 GTCTCCCTTCTTTCTTCACCTTATCTTAATGATC......TCCTGCC 793

233 lnArgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeu 249
    :::::::::|:|||:::||||
794 CTCAGAACGTCTACTGCTCTTGCTTTATG.............. 822

250 PheTyrGlyThrAlaIleIleIleMetTyrValGlyProArgTyrGlyAsnPr 266
    ||| |||
823 ...........TCTTACTTCAAC. 834

266 oLysGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProM 283
    |||||||:::||:|:
835 ...........ATGTACCTTATACTCATC...... 865

283 etLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThr 299
    ::|||||||||:::||||||::::::||||||::::::||
866 TGATCGATCCTCTCATCTACGCCCTCCGCAGCCCAAGAGATGCGGAGGACC 915

300 LeuLysArgValLeu 304
    :::|||:::::::
916 TTTAAGGAGATCGTC 930
```

Left column (sequence alignment)

```
  84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr 100
 680 TCCTGCATCCAGCCAAGCCCATCTCTTTGCTGGTTGCATGACGCAGACC 631

 101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
 630 TTTCTCGTTTGAGTTTGGACAGAGGAAATGTCTCCTGCTGGTGCTGAT 581

 117 tSerTyrTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
 580 GTCCTACGATCGTTACGTGGCCCAATCTGCCACCCTCCGAATACTCCCTCA 531

 134 leMetThrThrArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
 530 TCATGACCACGGGAGAGTGCATCACCCTGGCCGTCACTTCCTGGACGTGT 481

 151 GlyValLeuSerLeuHisLeuValLeuLeuLeuProLeuProh 167
 480 GGCTCCTCCTGCTGTTCCCCAGTGGTTCCTCATCCTAAGACTGCCCTT 431

 167 eCysArgProGlnLysIleTyrHisPheCysGluIleLeuLeuAlaValt 184
 430 CTCTCGGCCCTCATGAAATCAACCACTTCTTCTTGTGAAATCCTGTTCTGTCC 381

 184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
 380 TCAGGCTGGCCTGTGCCTGACACCTGGTCTACACCATCTTTGCA 331

 201 GlyValAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
 330 GGCCGTCGCTGTTCTTCCTGGTGGGGCCCACCAGCCTGTGCTTGTCTCCTA 281

 217 rMetCysIleLeuLeuAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
 280 CTCGCACATCCTGGCGCCCATCCTGAGGATCCAGTCTCGGGAGGGCCGCA 231

 234 rgLysAlaPheCysSerPheSerHisLeuCysValIleGlyLeuPhe 250
 230 GAAAGGCCTTCTCCACCGTCTCCCACCTCTGCGTGGTGGGACTCTTC 181

 251 TyrGlyThrAlaIleIleMetThrYrValGlyIleValGlyAsnProLy 267
 180 TTTGGCAGTCCAATCACATCANTGTACATGGCCCCCAAGTCCCGCCATCCTGA 131

 267 sGluGlnLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetLt 284
 130 GGAGCAGCAAAAGGTGTTTTTTCTATTTCACGTTTTTTCAACCCAACAC 81

 284 euAspProLeuLeuCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
  80 TTAACCCCCTGATTTACACCCTGAGGAACGGAGGAGTTCAAGGGTGCCCTG 31

 301 LysArgValLeuGlyValGlu 307
  30 AGGAGAGCACTGGGCAAGGAA 10
```

Documentation block

alignment_block:
US-09-898-586-24 x AAI23185

Align seq 1/1   to reverse of: AAI23185   from: 1   to: 933

alignment_scores:
```
        Quality:   1126.00        Length:    307
          Ratio:     4.155          Gaps:      0
Percent Similarity:  88.274   Percent Identity:  71.336
```

```
   1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
     |||||:|:|||||   ||| ::|||||||||||||||||||||||||||||
   1 ATGGGAGAAAATCAGACACAATGTGCACAGAGTTCCTCCTACTGGGATTT 50

  17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
     ::||||||||||||||||::||||||||||||||||||:|||||||||||
  51 CCTGGGCCCAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTTCT 100

  34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
     ||:::||||||||||||||:||||:|||||||||||||||||||||||||
 101 ATATCTTCACCCTGCTGGGAAACGGGACCATCCTGGGCCTCATCTCACTG 150

  51 AspSerArgLeuHisLeuAlaProMetTyrPheLeuLeuSerHisLeuAla 67
     |||||||||||:::||||||||||||||||||||||||||||||||||||||
 151 GACTCCAGACTCCACACCCCCAATGTACTTCCTCCTCTCCCACACCTGGTGT 731
```

```
  67 AlaSerArgLeuHisLeuAlaProMetTyrPheLeuLeuSerHisLeuAla 67
 780 GACTCCAGACTCCACACCCCAATGTACTTCCTCCTCTCCCACACCTGGTGT 731
```

```
  84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr 100
 680 TCCTGCATCCAGCCAAGCCCATCTCTTTGCTGGTTGCATGACGCAGACC 631

 101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseq-embl/NA2001A.DAT:ABA68294

seq_documentation_block:
ID   ABA68294 standard; DNA; 933 BP.
XX
AC   ABA68294;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16599.

XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-2074456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632256.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver -
PS   Claim 4; SEQ ID NO 16599; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 933 BP; 243 A; 220 C; 303 G; 167 T; 0 other;

alignment_scores:
       Quality: 1126.00   Length: 307
         Ratio: 4.155   Gaps: 0
   Percent Similarity: 88.274   Percent Identity: 71.336

alignment_block:
   US-09-898-586-24 x ABA68294/rev

Align seg 1/1 to reverse of: ABA68294   from: 1   to: 933

```
780 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCCTCAGCCACCTGGCTGT
    |||||||||||||||:::||||||||||||||||||||||||||||||
 67 IValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
730 CGTCGACATCGCCTACGCCTGCAACACCGTGCCCAGATGCTGGCGAACC 681

 84 euLeuHisProAlaLysProIleSerPhePheAlaGlyTyrArgMetMetGlnThr 100
680 TCCTGCAATCCAGCCAAGCCCATCTCCTTTGCTGGTTGCATGAGGCAGACC 631

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
630 TTTCTCTGTTTCAGTTTGGACACAGCAAGTCTCCTGCTGGTTGGCTGAT 581

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
580 GTCCTACGATCTTACGTGGCCATCTGCCACCCTCTCCGATACTCCGTCA 531

134 leMetThrThrArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
530 TCATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACTTCCTGGACGTCT 481

151 GlyValLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProH 167
480 GGCTCCCTCTCCCTGGCTCTGGCTGTCTCATCCTAAGACTGCCCTT 431

167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
430 CTCTGGGCCTCATGAAAATCAACACCTTCTTCTGTGAAATCCTGTCTGTCC 381

184 euIleLysAlaCysSerAspThrHisIleAsnGluAsnMetValLeuAla 200
380 TCAGGCCTGGCCTGTGCTCAACACCTGCCTCAACCAGGTGGTCATCTTTGCA 331

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleLeuValSerTy 217
330 GCCTGCGGTCGTCTTCCTGGTGGGGCCCACCCAGCCTCCGGTCGTGTCCTTA 281

217 rMetCysIleLeuLeuCysSerAlaIleLeuGlnIleSerArgArgGlyLeuValGlnA 234
280 CTCGCACACATCCTGGGCGCCATCCTGAGACAGTCTGGGGAGGGCCGGCA 231

234 rgLysAlaPheCysSerThrCysPheSerHisLeuCysSerValIleGlyLeuLeuGly 250
230 GAAAGGCCTTCTCCACCTGCTCCTCCCGCTTTCTATTTACAGTTTTTTCAACCCAAC 181

251 TyrGlyThrThrAlaIleIleLeuMetThrTyrValGlyProArgTyrGlyAsnProV 267
180 TTTGGCAGTGCCATCAACATCAGTAGACAATGGACTCCCGACATCCTGA 131

267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheLeuAsnProMetL 284
130 GGAGCAGCAAAAGGTCTTTTTTCTATTTACAGATTTTCAACCCAAC 81

284 euAsnProLeuLeuCysSerSerLeuArgAsnSerGluValLysAsnThrLeu 300
 80 TTAACCCCTGATTACAGCCTGAGGAACGGAGAGAGTCAAGGGTGCCCTG 31

301 LysArgValLeuGlyValGlu 307
 30 AGGAGAGGCACTGGGCAAGGAA 10
```

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
    |||||||||||||||||||||||||||||||||||||||||||((( 
  1 ATGGGGGAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 881

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
    ::|||||||||||||||||||||||||||||||||||||||||||||||
880 CCTGGGCCCAAGGAATTCAGATGCTCCTCTTTGGGCTCTTCCCCGTTTCT 831

 34 yrValPheThrLeuLeuGlyLeuAsnGlyThrLeuLeuGlyIleLeuSerLeu 50
    |||::|||||||||||||||||:::||||||||||||||||||||||||
830 ATATCTTCACCTCTGGGAGAACGGGCGCCATCCTGGGCGGTCCATCTCACTG 781

 51 AspSerArgIleHisLeuAlaProMetTyrPhePheLeuLeuSerHisLeuAlaVa 67
    ||:|||||||||||::||||||||||||||||||||||||||||||||||||
780 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCCTCAGCCACCTGGCTGT 731

 67 IValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
    |||||||||||||||:::|||||||||||||||||||:::||||||||||
730 CGTCGACATCGCCTACGCCTGCAACACCGTGCCCAGATGCTGGCGAACC 681
```

```
201  |||:||||||||||||||||||||||||||||||||||||||||||||:::
     CGTCGACACGCCTACACCCCGACAACACGGTGCCCAAGAATGCTGGCGAACC  250

 84  euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr
     |||||||||||||||||||||||||||||||||||||||||||  |||
251  TCCTGCATCCAGCCAAGCCCAATCCTTCCTTCGCTGGCCGCCATGACGCAGACC  300

101  PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe  117
     |||||    |||||||:  ||||::|||||||||:||||||||||||||||
301  TTTCTGTTCTTTGTTTGAGTTTTGGACAGCGAAGTCTCTGCTGGTGCTGAT  350

117  tSerTyrArgAspLeuThrValAlaIleAlaCysHisSerProLeuArgTyrLeuAla  134
     |||||||    |||||||||||||||||||||||||||||||||  ::|
351  GTCCTACGATCGTTACGTCGCCATCGCCTGCCACTCCCCGATACTCCGTCA  400

134  leMetThrThrProArgValCysIleThrLeuAlaValThrThrSerTrpThrThr  150
     |||||||||||||||||||||||||||||||||||||||||||||||||
401  TCATGACCACTCGGAGAGTGTGCATCACCCTGGCCGTCACTACTTCCTGGACCTGT  450

151  GlyValLeuLeuSerIleLeuHisIleLeuValLeuLeuLeuProLeuProh  167
     ||||||||||||||||||::|||                      ||||||
451  GGCTCCTCCTCTGGGCTCGGCCCATGTGGTTTCTCATCCTAAGACTGCCCTT  500

167  eCysArgProGlnLysIleTyrHisPheCysSGluIleLeuAlaVall  184
     |||::||||||||||||||:|||||||||||||||||||::||
501  CTCTGGGCCTTCATCAAATACAACCACTCTTCTGTGAAATCCTGTCTGTCC  550

184  euLysLeuLeuAlaCysAlaAspThrHisIleLeuAsnGluAsnMetValLeuAla  200
     |::||||||||||||||||||||||||||||||||||||||||||||||
551  TCAGGCTGGCCTGTGTCTGACACCTGGGCCATCAACTCACTCTTTGCA  600

201  GlyAlaIleLeuSerGlyLeuValIleGlyProLeuSerThrIleValValSerTy  217
     ||||||||||||||||||||||||||||||||||||||||||||::||||
601  GCCTGCGGTGTTCTTCCTGGTGGGGCCACCCAGCCTGTGTCTTGTCTCCTA  650

217  rMetCysIleLeuCysAlaIleIleLeuGlnIleGlnSerArgGluValGlnGln  234
     |||||||||||||||||||||||||||||||||||||||||||
651  CTCGGACATCCTGGCGGCCATCATCCTGCAGGGGATCCAGTCTCGGGAGGGGTCCGCA  700

234  rglLysAlaPheCysThrCysPheSerHisLeuCysSerValIleGlyLeuPhe  250
     ||||||||||||||||||||||||||||||:|||||||||
701  GAAAGGCCTTCTCCACCTGCTCTCCACCTGCGGTGGGGGGACTCTTC  750

251  TyrGlyThrAlaIleIleMetThrTyrValGlyProArgTyrGlyAsnProly  267
     :::|||:|||||||||||||||||||||||||||||||||||
751  TTTGGCAGTCCCATCATCATGACATGGCCCCCAAGTCCCGCCATCCTGA  800

267  sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetL  284
     :|||||||||||||||||||||||||||||||||||||||||||  |
801  GGAGCAGCAAAAGTCTTTTTTTCTATTTTACAGTTTTCAACCAACAC  850

284  euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu  300
     ||||||||||||||||||||||||||||::||||::|||||||||||
851  TTAACCCCTGATTTACAGCCTGAGGAACGGAGGTCAAGGGTGCCCTG  900

301  LysArgValLeuGlyValGlu  307
     :::|||||||:|||||  |||
901  AGGAGGACACTGGGCGAAGGAA  921
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:ABA50343

seq_documentation_block:
ID   ABA50343 standard; DNA; 933 BP.
XX
AC   ABA50343;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #9038.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207455.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PI   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PS   Claim 4; SEQ ID NO 9038; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 933 BP; 243 A; 220 C; 303 G; 167 T; 0 other;

alignment_scores:
           Quality: 1126.00   Length: 307
           Ratio: 4.155   Gaps: 0
Percent Similarity: 88.274   Percent Identity: 71.336

alignment_block:
US-09-898-586-24 x ABA50343/rev   ...

Align seg 1/1 to reverse of: ABA50343 from: 1 to: 933

```
  1  MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
     |||||:|||   |||  :::||||||||||||||||||||:||||||||||
930  ATGGGGGAAAATCAGACAATGGTCACAGAGTTCTCCTACTGGGATTTCT  881

 17  oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT  34
     :::|||||||||||||||||||:|||||||||||||||||||||||||||
880  CCTGGGGCCCAAGGAATCAGATGCTCCTCTTGGGCTCTTTCCCCGTTTCT  831

 34  yrValPheThrLeuLeuLeuGlyAsnGlyThrIleLeuIleLeuIleSerLeu  50
     ::|||||||||||||||||||::||||||||||||||::|||||||||||:
830  ATATCTTCACCCTCGCTGGGAACGGGCCCATCCTGGGGCTCAATCTCACTG  781

 51  AspSerArgLeuHisLeuAlaProMetThrTyrPhePheLeuSerHisLeuAlaVala  67
```

XX
XX Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 553; 1857pp; English.
XX
CC The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 930 BP; 165 A; 304 C; 219 G; 242 T; 0 other;

alignment_scores:
        Quality: 1126.00   Length: 307
          Ratio: 4.155   Gaps: 0
Percent Similarity: 88.274   Percent Identity: 71.336

alignment_block:
US-09-898-586-24 x AAH32363

Align seq 1/1 to: AAH32363 from: 1 to: 930

  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
     ||||| |||:|||||:|:||
  1 ATGGGGGAAAATCAGACAACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT 50

 17 oValGlyProArgIleGlnIleMetLeuPheGlyLeuPheSerLeuPheT 34
     :::||||||||||||||:|||||||||||||||||||:||||||
 51 CCTGGGCCCAAGGATTCAGATGCTCTTCTTTGGGCTCTTCTCCCTGTTTCT 100

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuLeuGlyLeuIleSerLeu 50
     |::||||||||||||||||||||||||||||||||||||
101 ATACTTCACCTGCTGGGAACAGGGCCATCCTGGGCTCATCTCCACTG 150

 51 AspSerArgLeuHisSerAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
     ||||||:|:||:|:|||||||||||||||||||||||||||||||||||
151 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCTCACACCTGGCTGT 200

 67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84

```
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
    |||||||||||||||::|||||||||:::||||||||||
885 TCATGACCTGGAAAGTCTGCATCACTCTGGCATCACTTCTTGGACATGT 934

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProh 167
    |||||||||||||:::|||||||||||||:::|||||||||||||
935 GGCTCCCTCCTGGCTATGGTCCATGTGAGCCCTCATCCTAAGACTGCCCTT 984

167 eCysArgProGlnLysIleTyrHisPheCysSGlurIleLeuAlaValL 184
    |||||    ||||||:::||:||||||||||||||||||||:::|||
985 TTGTGGGGCCTTCGTGAAAATCAACCACTTCTTCTTCTGTGAAATCCTGTCTCTCC 1034

184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
    |||||||||||||||||||||||||||||||||||||||||||
1035 TCAGGCTGGCCTGTGCTGATACCCTGGTCCACCAGGTGGTCATCTTTGCA 1084

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
    :::::||||||||||||||||||:::||||||||||||||||
1085 GGCTGCCGCATGTCATCCGTGGGGACCCACTCTCCTGGTGGTCGTCCCTA 1134

217 rMetCysIleLeuLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
    |||||||||||||||:::||||||||||||||||||||:::|
1135 CTCACACATCCTGCGCGCCCCATCCTGAGGATCCAGTCTCGGGAGGGCCGCA 1184

234 rgLysAlaPheCysTHrCysPheSerHisLeuCysValIleGlyLeuPhe 250
    ||||||||||    |||||||||||||::|||||||||||||
1185 GAAAGGCCTTCTCCACCGTCTCTCCACCCTCGCGTAGTCGGACTCTTC 1234

251 TyrGlyThrAlaIleIleLeuMetTyrValGlyProArgTyrTyrGlyLyAsnProLy 267
    ::||||||||||||||:::|::|||||
1235 TTTGGCAGCCACGTCATGTACAGTGCCCCCTAAGTCCCGCCATCCTGA 1284

267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
    ::|||:||:|||||||||||||||:::||||||||||||||
1285 GGAGCAGAAGAAGTCCTTTTTCTCTATTTACAGTTCTTCAACCCGATGC 1334

284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
    ||||||||||||||||||||||||||    |||||||||||||||
1335 TAAACCCCTGATTTACAACCTGAGGAATGTAGAGGTCAAGGGTGCCCTG 1384

301 LysArgValLeu 304
    :::|||:::||
1385 AGGAGAGCACTG 1396
```

seq_documentation_block:
```
ID    AAS08654 standard; cDNA; 2735 BP.
XX
seq_name: /SIDS1/gcgdata/hold-geneseqn/geneseqn-embl/NA2001A.DAT:AAS08654

DT    17-DEC-2001 (first entry)
XX
DE    Human cDNA encoding partial GPCR-like protein, GPCR #9.
XX
KW    Human; G-protein coupled receptor like protein; GPCR; immunogen;
KW    ophthalmic disease; neurological disease; Alzheimer's disease;
KW    Parkinson's disease; immunological disorder; HIV; candidiasis;
KW    human immunodeficiency virus; autoimmune disorder; multiple sclerosis;
KW    systemic lupus erythematosus; rheumatoid arthritis; platelet disorder;
KW    septic shock; systemic inflammatory response syndrome; SIRS;
KW    thrombocytopaenia; aplastic anaemia; inflammatory disorder;
KW    hormonal dysfunction; cancer; atherosclerosis; wound;
KW    tissue regeneration; haemophilia; leukaemia; reperfusion injury;
KW    psoriasis; diabetes; ss.
XX
OS    Homo sapiens.
XX
FH    Key            Location/Qualifiers
FT    CDS            1206..2267
FT                   /tag= a
FT                   /product= "GPCR #9"
```

```
FT                   /partial
FT                   note= "No start codon"
XX
PN    WO200153454-A2.
XX
PD    26-JUL-2001.
XX
PF    22-DEC-2000; 2000WO-US34983.
XX
PR    21-JAN-2000; 2000US-0488725.
PR    25-APR-2000; 2000US-0552317.
PR    20-JUN-2000; 2000US-0598042.
PR    19-JUL-2000; 2000US-0620312.
PR    31-AUG-2000; 2000US-0653450.
PR    04-DEC-2000; 2000US-0729739.
XX
PA    (HYSE-) HYSEQ INC.
XX
PI    Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F;
PI    Asundi V, Drmanac RT;
XX
DR    WPI; 2001-442255/47.
DR    P-PSDB; AAU04580.
XX
PT    New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT    useful for treating diseases of ophthalmic, neurological, immunological
PT    and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT    and diabetes.
XX
PS    Claim 1; Page 212-214; 259pp; English.
XX
CC    The sequence encodes a human G-protein coupled receptor (GPCR)-like
CC    protein. The GPCR-like polypeptides and polynucleotides are useful for
CC    the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's
CC    disease and Parkinson's disease), immunological (e.g HIV infection and
CC    candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic
CC    lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g.
CC    thrombocytopenia and aplastic anaemia), inflammatory disorders (e.g.
CC    septic shock and systemic inflammatory response syndrome, SIRS) and
CC    nephritic systems. They may also be used to treat hormonal dysfunction,
CC    cancer, atherosclerosis, wounds, tissue regeneration, haemophilia,
CC    leukaemias, reperfusion injury, psoriasis and diabetes. Numerous examples
CC    of each type of disorder are given in the specification. Anti-GPCR-like
CC    protein antibodies are useful for detecting or quantitating the
CC    polypeptide in tissue. The polypeptides can also be used as molecular
CC    weight markers and as a food supplement.
XX
SQ    Sequence 2735 BP; 695 A; 671 C; 520 G; 849 T; 0 other;
```

alignment_scores:
```
          Quality: 1138.00          Length:    304
          Ratio:  4.246             Gaps:      0
Percent Similarity: 88.158     Percent Identity: 72.039
```

alignment_block:
```
Align seg 1/1  to: AAS08654
US-09-898-586-24 x AAS08654

Align seg 1/1  to: AAS08654  from: 1  to: 2735
```

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
    |||   |||              |||    ||||||||||||||||||
1335 ATGGTGAAAAATCAGACAATGGTCCACACAGAGTTCCTCCTACTGGGATTTCT 1384

 17 ovalGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPheT 34
    ::||||||||||||||||||||||||||||||||||||||||||
1385 CCTGGGCCCAAGGAGATTCAGATGCTCCCCTTTGGGCTCTTCTCCCTGTTTCT 1434

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
    |||||||||||||::|||||||||||||||::|||||||||||||||
1435 ATGTCTTCACCCTCGTCGGGAATGGGACCATCCTGGGGCTCATCTCACTG 1484

 51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
```

```
1907 CTCACACATCCTGCGCGGCGGCAATCCTGAGGAATCCAGTCTGGGGAGGCGGCCCA  1956
          ||||||:::||||||||:|||||||||||||||||||:||||||||||||   ::||
1957 GAAAGGCGCTTCTCCACCTTGCTCCTCCACCTGGTCCACCTTCTGGTAGTGGGGACTCTTC  2006

234  rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe  250
          ||||||:||||| ||||||||||||||||||||||||||||||||
1957 GAAAGGCGCTTCTCCACCTTGCTCCTCCACCTGGTCCACCTTCTGGTAGTGGGGACTCTTC  2006

251  TyrGlyThrAlaIleIleMetTyrrValGlyProArgTyrrGlyIyAsnProly  267
          ||:||||||||||||||||||||||||||||||||||||||:::
2007 TTTGGCAGCGCCATCATCATGTACATGGCCCCTAAGTCCCGCCATCCTGA  2056

267  sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetL  284
          ||||::|||||||||||||||||||||||||||||||||||||||||
2057 GGAGCAACGAGAAGTACCTTTTCTTTTCATTTTCATTTCTTCAACCGGATGC  2106

284  euAsnProLeuIleCysSerLeuArgAsnSerGluVallysAsnThrLeu  300
          ||||||||| |||||||||||||::|||||||||:::||||
2107 TAAACCCCTGATTTACAAACCTGAGGAAATGTAGGAGTCAAGGGTGCCCTG  2156

301  LysArgValLeu  304
          ::||:|||||||
2157 AGGAGCGACTG  2168
```

```
seq_name:  /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS08655

seq_documentation_block:
ID    AAS08655 standard; cDNA; 1788 BP.
AC    AAS08655;
XX
DT    17-DEC-2001 (first entry)
XX
DE    Human cDNA encoding G-protein coupled receptor like protein, GPCR #10.
XX
KW    Human; G-protein coupled receptor like protein; GPCR; immunogen;
KW    ophthalmic disease; neurological disease; Alzheimer's disease;
KW    Parkinson's disease; immunological disorder; HIV; candidiasis;
KW    human immunodeficiency virus; autoimmune disorder; multiple sclerosis;
KW    systemic lupus erythematosus; rheumatoid arthritis; platelet disorder;
KW    thrombocytopaenia; aplastic anaemia; inflammatory disorder;
KW    septic shock; systemic inflammatory response syndrome; SIRS;
KW    hormonal dysfunction; cancer; atherosclerosis; wound;
KW    tissue regeneration; haemophilia; leukaemia; reperfusion injury;
KW    psoriasis; diabetes; ss.
XX
OS    Homo sapiens.
XX
FH    Key                 Location/Qualifiers
FT    CDS                 833..1417
FT                        /*tag= a
FT                        /product= "GPCR #10"
FT                        /note= "This sequence is specifically claimed"
FT    sig_peptide         833..937
FT                        /*tag= b
FT    mat_peptide         938..1414
FT                        /*tag= c
FT                        /label= Mature_GPCR #10
XX
PN    WO20015454-A2.
XX
PD    26-JUL-2001.
XX
PF    22-DEC-2000; 2000WO-US34983.
XX
PR    21-JAN-2000; 2000US-0488725.
PR    25-APR-2000; 2000US-0552317.
PR    20-JUN-2000; 2000US-0598042.
PR    19-JUL-2000; 2000US-0620312.
PR    31-AUG-2000; 2000US-0653450.
PR    04-DEC-2000; 2000US-0729739.
XX
PA    (HYSE-) HYSEQ INC.
```

```
PI    Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F;
PI    Asundi V, Drmanac RT;
XX
DR    WPI; 2001-442255/47.
DR    P-PSDB; AAO04574.
XX
PT    New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT    useful for treating diseases of ophthalmic, neurological, immunological
PT    and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT    and diabetes -
XX
PS    Claim 1; Page 214-215; 259pp; English.
XX
CC    The sequence encodes a human G-protein coupled receptor (GPCR)-like
CC    protein. The GPCR-like polypeptides and polynucleotides are useful for
CC    the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's
CC    disease and Parkinson's disease, immunological (e.g HIV infection and
CC    candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic
CC    lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g
CC    thrombocytopaenia and aplastic anaemia), inflammatory disorders (e.g.
CC    septic shock and systemic inflammatory response syndrome, SIRS) and
CC    nephritic systems. They may also be used to treat hormonal dysfunction,
CC    cancer, atherosclerosis, wounds, tissue regeneration, haemophilia,
CC    leukaemias, reperfusion injury, psoriasis and diabetes. Numerous examples
CC    of each type of disorder are given in the specification. Anti-GPCR-like
CC    protein antibodies are useful for detecting or quantitating the
CC    polypeptide in tissue. The polypeptides can also be used as molecular
CC    weight markers and as a food supplement.
XX
SQ    Sequence 1788 BP; 436 A; 446 C; 342 G; 564 T; 0 other;

alignment_scores:
              Quality: 1138.00          Length:    304
              Ratio:      4.246         Gaps:        0
Percent Similarity:  88.158    Percent Identity: 72.039

alignment_block:
US-09-898-586-24 x AAS08655

Align seq 1/1  to: AAS08655        from: 1  to: 1788

    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
          ||| |:||:||||||||:::||||||||||||||||||||||||||||
  485 ATGGTGAAAAATCAGACAATGGTCACAGAGTTCCTCCTACTGGGATTTCT  534

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheHer  34
          ::||||||||||||||||||||||||:||||||||||||||||||||||
  535 CCTGGGCCCAAGGATTCAGATGCTCCTCCTTGGGCGTCTTCTCCCGTTCT  584

   34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu  50
          |||||||||||||||||||||||||||||||||||||||||||| |||
  585 ATGTCTTCACCCTGCTGGGAAAGGGACCAATCCTGGGCTCATCTCACTG  634

   51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa  67
          ||||||||||||::|||:||||||||||||||||||||||||||||
  635 GACTCCAGACTCCACACCCCCAAGTACTTCTTCCTCTCACCTGCCGCGT  684

   67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL  84
          ||| |:|||||||||||||||||||||||||||||||||||||||||
  685 CGTCAACATCGCCTATGCCTGCAACACAGTGCCCCAGATGCTGGTGAACL  734

   84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr  100
          |||||||||||||||||||||||||||||||||||||      |||
  735 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGCCATGACATAGACC  784

  101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe  117
          ||||||||||||||:||||||||||||||||||||||||||||||||
  785 TTTCTCTTTTTGAGTTTTGCACATACCTGAATGCCTCCTGTTGGTGCTGAT  834

  117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI  134
          |||||||||||||||||||||||||||||||||||||||||||:::
  835 GTCCTACGATCGGTACGTGGCCATCTGCCACCCCTCCTCGATATTTCATCA  884
```

```
         //////||||||||
      85 GCTCAATCCCGTTATCTGTAGTCGTTAGGAACTCAGAAGTGAAGAATACTT 36
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     300 eulysArgValLeuGlyValGluArgAlaLeu 310
         |||||||||||||||||||||||||||||||
      35 TGAAGAGAAGTGTGGGAGGTAGAAAAGGGCTTTA 4
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS08660

seq_documentation_block:
ID   AAS08660 standard; cDNA; 2282 BP.
XX
AC   AAS08660;
XX
DT   17-DEC-2001 (first entry)
XX
DE   Human cDNA encoding G-protein coupled receptor like protein, GPCR #13.
XX
KW   Human; G-protein coupled receptor like protein; GPCR; immunogen;
KW   ophthalmic disease; neurological disease; Alzheimer's disease;
KW   Parkinson's disease; immunological disorder; HIV; candidiasis;
KW   human immunodeficiency virus; autoimmune disorder; multiple sclerosis;
KW   systemic lupus erythematosus; rheumatoid arthritis; platelet disorder;
KW   thrombocytopaenia; aplastic anaemia; inflammatory disorder;
KW   septic shock; systemic inflammatory response syndrome; SIRS;
KW   hormonal dysfunction; cancer; atherosclerosis; wound;
KW   tissue regeneration; haemophilia; leukaemia; reperfusion injury;
KW   psoriasis; diabetes; ss.
XX
OS   Homo sapiens.
XX
FH
FT   Key   Location/Qualifiers
FT   CDS   1257..2189
FT   /tag= a
FT   /product= "GPCR #13"
XX
PN   WO200153454-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US34983.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   20-JUN-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   31-AUG-2000; 2000US-0653450.
PR   04-DEC-2000; 2000US-0729739.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F;
PI   Asundi V, Drmanac RT;
XX
DR   WPI; 2001-442255/47.
DR   P-PSDB; AAU04582.
XX
PT   New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT   useful for treating diseases of ophthalmic, neurological, immunological
PT   and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT   and diabetes.
XX
PS   Claim 1; Page 256-258; 259pp; English.
XX
CC   The sequence encodes a human G-protein coupled receptor (GPCR)-like
CC   protein. The GPCR-like polypeptides and polynucleotides are useful for
CC   the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's
CC   disease and Parkinson's disease, immunological (e.g HIV infection and
CC   candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic
CC   lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g
CC   thrombocytopaenia and aplastic anaemia), inflammatory disorders (e.g
CC   septic shock and systemic inflammatory response syndrome, SIRS) and
CC   nephritic systems. They may also be used to treat hormonal dysfunction,
CC   cancer, atherosclerosis, wounds, tissue regeneration, haemophilia,
CC   leukaemias, reperfusion injury, psoriasis and diabetes. Numerous examples
CC   of each type of disorder are given in the specification. Anti-GPCR-like
CC   protein antibodies are useful for detecting or quantifying the
CC   polypeptide in tissue. The polypeptides can also be used as molecular
CC   weight markers and as a food supplement.
XX
SQ   Sequence 2282 BP; 583 A; 573 C; 428 G; 698 T; 0 other;

alignment_block:
us-09-898-586-24 x AAS08660

Align seg 1/1   to: AAS08660   from: 1   to: 2282

alignment_scores:
     Quality: 1147.00   Length: 304
     Ratio: 4.264   Gaps: 0
     Percent Similarity: 88.487 Percent Identity: 72.368
```

```
    1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
      |||    :::||||||||||||||||||||||||||||||||||||||
 1257 ATGGTGAAAAATCAGACAAATGGTCACAGAGAGTTCCTTCCTACTGGGATTTCT 1306

   17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheI 34
      ::||||||||||||||||||||||||||||||||||||||||||:|||
 1307 CCTGGGCCCAAGGATTCAGATGCTCCTTCTTTGGGCTCTTCTCCCGTTCT 1356

   34 yrValPheThrLeuGlyAsnGlyThrIleLeuLeuGlyLeuIleSerLeu 50
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1357 ATGTCTTCACCCTGCTGGGGAATGGGACCATCCTGCTGGGGCTCATCTCACTG 1406

   51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
      ||||||||:::||||||||||||||||:|||||||||||||||||||||
 1407 GACTCCAGACTCCACACCCCATGTACTTCTTCCTCTCACACCTGGCCGT 1456

   67 lValAspIleAlaTyrAlaCysSerAsnThrValProArgMetLeuValAsnL 84
      |||||||||||||||||||||||:||||||||||||||||||:||||
 1457 CGTCAACATCGCCTATGCCTGCAACACAGTGCCCCGAGATGCTGGTGAACC 1506

   84 euLeuHisProAlaLysProIleSerPheAlaGlyAlaGlyMetMetGlnThr 100
      |||||||||||:|||||||:|||||||||||||||||||||||||||||
 1507 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCTGGCGCTGGCATGAGACAGACC 1556

  101 PheLeuPheSerThrPheAlaValThrGlyCysLeuLeuLeuValAlaMe 117
      ||||||||:|||||||||||||||||||||||||||||||||||:|||
 1557 TTTCTCTTTTCGAGTTTTGCAGATACTGGATGCCTCCGTTGGGGTGCTGAT 1606

  117 tSerThrTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
      |||||||||:|||||||||||||||||||||||||||||||||::|
 1607 GTCCTACGATCGGTACGTGGCCATCTGCCACCCTCTCCGATATTCATCA 1656

  134 leMetThrTrpArgValCysIleLeuThrLeuAlaValThrSerTrpPheThr 150
      |||||||:::||||||||||||||||||||||||||||||::|||||
 1657 TCATGACCTGGAGAGTCTGCATCACTCTGGCCGTCACCTCCTGGACATGT 1706

  151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167
      |||||||||:::||||:||||::|||||||:|||||||||:||||||
 1707 GGCTCCCTTCCTGGCTAAGGTCCATAAGCTTGAGCCCATCCTAAGACTGCCCTT 1756

  167 eCysSerArgProGlnLysIleTyrHisSerPheCysGluIleLeuAlaValI 184
      ||||||||:::||||||||:::||||||||||||||||||||:||||
 1757 TTGTGGGCCCTCGTGAAATCAACCACTTCTTCTGTGAAATCCTGTCTGTCC 1806

  184 eulysLeuAlaCysSerAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
      |::::|||||||||||||||:::||||||::|||||||||::|||:|||
 1807 TCAGGCCTGCCCTGTCGTGCTGATACCTGCCAACCAGTGGCTCATCTTTGCA 1856

  201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
      ||||||||||||||||||||||||||||||||:||||||||||||||
 1857 GCCTGCAATGTTCAATCCCTGGTGGGACCACTCTCCCCTGTGCTGGTCTCCTA 1906

  217 rMetCysIleLeuCysAlaIleLeuLeuGlnIleGlnSerArgGluValGlnA 234
```

```
KW  antifungal; vulnerary; antiulcer; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157188-A2.
XX
PD  09-AUG-2001.
XX
PF  05-FEB-2001; 2001WO-US03800.
XX
PR  03-FEB-2000; 2000US-0496914.
PR  27-APR-2000; 2000US-0560875.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Tang YT, Liu C, Drmanac RT;
XX
PT  Human proteins and DNA encoding sequences useful for preventing,
PT  treating or ameliorating a medical condition in a mammalian subject
PT  e.g. arthritis and cancer -
XX
PR  WPI; 2001-457740/49.
PR  P-PSDB; ABB11829.
XX
PS  Claim 1; Page 753; 1963pp; English.
XX
CC  Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC  sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC  invention also relates to vectors and recombinant host cells comprising a
CC  nucleotide of the invention, methods of producing the novel polypeptides,
CC  antibodies against the polypeptides, methods of identifying compounds which
CC  bind to polypeptides in a sample, and methods of detecting the nucleotides
CC  or polypeptides of the invention. Although novel, many of the
CC  polypeptides of the invention have homology to known proteins, thereby
CC  giving an insight into their probable biological activities, and hence
CC  potential therapeutic applications. The polypeptides of the invention may
CC  have various activities, including cytokine, cell proliferation or cell
CC  differentiation activities; stem cell growth factor activity;
CC  haematopoiesis regulatory activity; tissue growth activity;
CC  immunomodulatory activity; activin- or inhibin-related activities;
CC  chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC  thrombolytic activities; receptor or ligand activities; or may be
CC  involved in oncogenesis, cancer cell proliferation or metastasis.
CC  Depending on their biological activities, polypeptides and nucleotides of
CC  the invention are useful for preventing, treating or ameliorating medical
CC  conditions, e.g., by protein or gene therapy. Such conditions include
CC  cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC  disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC  proliferative retinopathy, atherosclerosis, coronary heart disease,
CC  vascular growth. Polypeptides involved with tissue regeneration and
CC  repair (or nucleic acids encoding them) may be used to promote wound
CC  healing (e.g., of burns, incisions and ulcers), while those with
CC  immunomodulatory activities may be used in the treatment of viral,
CC  bacterial and fungal infections in addition to immune disorders.
CC  Polypeptides with growth factor activity may be used in cell cultures to
CC  promote cell growth. For example, such polypeptides may be used to
CC  manipulate stem cells in culture to give rise to neuroepithelial cells
CC  that can be used to augment or replace cells damaged by illness,
CC  autoimmune disease or accidental damage. The polypeptides and nucleotides
CC  may also be used in the diagnosis of the above conditions, and in drug
CC  screening techniques. The present sequence represents a cDNA encoding a
CC  novel human polypeptide of the invention.
XX
SQ  Sequence 1315 BP; 380 A; 281 C; 360 G; 294 T; 0 other;
```

```
alignment_scores:
       Quality:   1583.00              Length:    312
         Ratio:      5.106               Gaps:      2
Percent Similarity:  99.359      Percent Identity: 99.359

alignment_block:
```

```
US-09-898-586-24 x ABA09073/rev  ..

Align seq 1/1  to reverse of: ABA09073  from: 1  to: 1315


     1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
       ATGGGGGACAATATAACATCCATCACAGAGTTCCTGCTACTGGGATTTCC 886

    17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
       CGTTGGCCCAAGGAATTCAGATGCTCCTTCTTTGGGCTCTTTCCTGTTCT 836

    34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
       ACGTCTTCACCCTGCTGGGGAACGGGACCATACTGGGGCTCATCTCACTG 786

    51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeu.AlaV 67
       GACTCCAGACTGCACGCCCCCATGTACTTCTTTCCTCTCACACCTGGCCG 736

    67 alValAlaSpIleAlaTyrAlaLysAsnThrValProLeuArgMetLeuAsn 83
       TCGTGGCAACATGGCCTACGCCTGCAAACACGGTGCCCCGGATGCTGGAAC 686

    84 LeuLeuHisProAlaLysProIleSerProPheAlaGlyAngMetMetGlyrh 100
       CTCCTGCATCCAGCCAAGCCCATCTCCTTGCAGGGCCAGATGATGCAGAC 636

   100 rPheLeuPheSerThrPheAlaValAlaThrGluCysLeuLeuLeuValMet 117
       CTTTCTCTTTCCACTTTTGCTGTCGCAACAGAGTGCCTCCTGGTGGTGA 586

   117 etSerThrTyrAspLeuLeuTyrVal.AlaIleCysHisProLeuArgTyrLeuA1 133
       TGTCCTTATGAGATCTGCAGTGGGCCCAATCTGCCACCCCTCCGGATATTTGGC 536

   133 alIleMetThrTrpArgValCysIleThrLeuAlaValThrServTrpThrT 150
       CATCATGACCTGGAGAGTCTGCATCACCCTGCGCGTGACTTCCTGGACA 486

   150 hrGlyValLeuLeuSerLeuHisLeuHisLeuValLeuLeuProLeuPro 166
       CTGGAGTCCTTTTATCCTTGAATTCCATTGTGTTACTTCTACCTTTACC 436

   167 PheCysArgProGlnLysIleTyrHisSerPhePheCysGluIleLeuAlaVa 183
       TTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTGTGAAATCTTGGCTGT 386

   183 lLeuLysLeuAlaLysAlaAspThrHisIleAsnGluAsnMetValLeuA 200
       TCTCAAACTTGCCTGTCCAGATACCCACATCAATGAGAACATGGTCTTGG 336

   200 laGlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSer 216
       CCGGAGCCAATTTCTGGGCTGGTGGGACCCCTGTCCACAATTGTAGTTTCA 286

   217 TyrMetCysIleLeuCysSerAlaIleLeuGlnIleSerArgGluValGl 233
       TATATGTGCATCCTCTGTGCTATCCTTCAGATCAGCAGGGAAGTTCA 236

   233 nArgLysAlaPheCysSerThrCysPheSerHisLeuCysValIleGlyLeuP 250
       AACGGAAGGCCTTCTGCAACCTCGTCTTCCCACCTCGTGTGATTGGACTCT 186

   250 heTyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnPro 266
       TTTATGGCACAGCCATTATCAATGTATGTTGGACCCAGATAGGGAACCCC 136

   267 LysGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMe 283
       AAGGAGCAGAAGAAATATCTCCTGCTGTTTCACAGCCTCTTTAATCCAT 86

   283 tLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrL 300
```

PR   24-JUL-2000; 2000US-0220253.
PR   25-JUL-2000; 2000US-0220590.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;
PI   Spytek KA, Li L;
XX
PI
DR   WPI; 2001-451859/48.
DR   P-PSDB; AAU05132.
XX
XX
PT   New NOVX polypeptides and polynucleotides, useful for treating or
PT   preventing disorders of the neuro-olfactory system, cancer and multiple
PT   sclerosis -
XX
PS   Claim 9; Page 15; 141pp; English.
XX
CC   The sequence represents the coding sequence of human odorant receptor
CC   (OR)-like protein, NOV2. The NOV2 polypeptide, nucleic acid and antibody
CC   are useful as therapeutics, particularly in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease, which
CC   includes a pathology associated with NOV2 polypeptide. The NOVX nucleic
CC   acid and polypeptide are especially useful in therapeutic or
CC   prophylactic applications for disorders of the neuro-olfactory system,
CC   e.g. those induced by trauma, surgery and/or neoplastic disorders. The
CC   DNA encoding the protein is useful in gene therapy for treating the
CC   above conditions. Furthermore, the nucleic acids and polypeptides are
CC   useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus
CC   cancer, immune response, acquired immunodeficiency syndrome (AIDS),
CC   asthma, Crohn's disease, multiple sclerosis or Albright hereditary
CC   osteodystrophy. These are also useful in developing powerful assay
CC   system for functional analysis of various human disorders, as well as
CC   in diagnostic applications.
XX
SQ   Sequence 1040 BP; 229 A; 285 C; 221 G; 305 T; 0 other;

alignment_scores:
          Quality: 1584.00   Length: 310
          Ratio: 5.160   Gaps: 0
Percent Similarity: 99.032   Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AAS09946

Align seg 1/1   to: AAS09946   from: 1 to: 1040
                AAS09946   from: 1 to: 1040

```
  1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 82 ATGGGAGACAATATAACATCCATCACAGAGAGTTCCTCTTACTGGGATTTCC 131

 17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
    |||||||||||||||||||||||||||||||||||||||||||||||||
132 CGTTGGCCCAAGGATTCAGAATGCTCCTCTTTGGCTCTTCTCCCTGTTCT 181

 34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuLeuGlyLeuIleSerLeu 50
    |||||||||||||||||||||||||||||||||||||||||||||||||||
182 ACGTCTTCACCCTGCTTGGGGCAACGGGACCATACTGGGGCCTCANCTCACTG 231

 51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
    |||||||||||||||||||||||||||||||||||||||||||||||||
232 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCCGGT 281

 67 lValAspIleAlaTyrValCysAsnThrValProArgMetLeuValAsnL 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
282 CGTTGACATTGCCTACGTCTGCAACACGGTGCCCCGAATGCTGGTGAACC 331

 84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
    |||||||||||||||||||||||||||||||||||||||||||||||||
332 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGATGCAGACC 381

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
    |||||||||||||||||||||||||||||||||||||||||||||||||
382 TTTCTGTTTTCCACTTTTGCTGTCACAGAAGTGTCTCCTCTGGGTGGTGAT 431

117 tSerTyrAspLeuThrTyrValAlaIleCysHisProLeuArgTyrLeuAla 134
    |||||||||||||||||||||||||||||||||||||||||||||||||||
432 GTCCTATGATCTCTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA 481

134 leMetThrTrpArgValCysIleThrLeuAlaValAlaThrSerTrpThrThr 150
    |||||||||||||||||||||||||||||||||||||||||||||||||||
482 TCATGACCTGGAGAGTGTGCATCACCCTCGCGGTGACTTCCTGGACCACT 531

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProh 167
    |||||||||||||||||||||||||||||||||||||||||||||||||
532 GGAGTCCTTTTACCTTGATTCATCTTGTGTTACTTCTACCTTTTACCCTT 581

167 eCysArgProGlnLysIleThrTyrHisProPhePheCysGluIleLeuAlaValL 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
582 CTGTAGGCCCCAGAAAATTTATCACTTTTTTTTTGTGAAATCTTGGCTGTTC 631

184 euLysLeuAlaCysAlaAspThrHisIleLeuAsnGluAsnMetValLeuAla 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
632 TCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 681

201 GlyAlaIleSerArgGlyLeuValAlyProLeuSerThrIleValValSerTy 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
682 GGAGCAAATTCTGGGCGTGGGACCCTTGTCCACAATTGTAGTTCATA 731

217 rMetCysIleLeuGlyCysAlaIleLeuGluGlnIleAspIleSerArgGluValGlnA 234
    |||||||||||||||||||||||||||||||||||||||||||||||||||||
732 TATGTGCATCCTCTGGTGTATCCTTCAGATCCAATCAAAGGAAGTTCAGA 781

234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
    |||||||||||||||||||||||||||||||||||||||||||||||||
782 GGAAAGCCTTCCGCACCTGCTTCTCCCACCTCTGTGTGATTGGACTCGTT 831

251 TyrGlyThrAlaIleLeuIleMetThrTyrValGlyProArgThrYrGlyAspProLy 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
832 TATGGCACAGCCATTATCATCATGATAGGAACCCCAA 881

267 sGlyLysLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
    |||||||||||||||||||||||||||||||||||||||||||||||||
882 GGAGGCAAGAAAATATCTCCTGCTGTTCACAGCCTCTTTAATCCCATGC 931

284 euAsnProLeuLeuCysSerLeuArgAsnSerGluValAllLysAsnThrLeu 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
932 TCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAACTCAGAATGAAGAAATACTTTG 981

301 LysArgValLeuGluGlyValGluArggAlaLeu 310
    |||||||||||||||||||||||||||||||||
982 AAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 1011
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:ABA09073

seq_documentation_block:
ID   ABA09073 standard; cDNA; 1315 BP.
XX
AC   ABA09073;
XX
DT   11-JAN-2002 (first entry)
XX
DE   Human olfactory receptor homologue-encoding cDNA, SEQ ID NO:849.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;

Claim 8: Page 264-265; 1857pp; English.

The present sequence is one of a number of isolated polynucleotides
which encode polypeptides involved in olfactory sensation. The
polynucleotides can be used in screening for olfactory agonists and
antagonists. The methods allow for the determination of primary
scents and the identification of the odour receptors used to detect
these primary scents. The methods also enable determination of
secondary scents and the identification of combinations of odour
receptors that are involved in detecting such secondary scents.
This enables the construction of a scent representation (also called
a scent fingerprint or scent profile), which may be used to re-create
and edit scents. Libraries of olfactory receptors are useful for
determining the interaction pattern of a composition with the receptors,
and can be used for determining differences in the olfactory faculties
of different individuals.

SQ   Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

alignment_scores:
     Quality: 1584.00
     Ratio: 5.160
     Percent Similarity: 99.032

alignment_block:
US-09-898-586-24 x AAH31648

Align seg 1/1 to: AAH31648
                                                             Length: 310
,   Gaps: 0
                 from: 1 to: 930   Percent Identity: 99.032

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
  ||||||||||||||||||||||||||||||||||||||||||||||||||
1 ATGGGAGACAATATAACATCCATCACAGAGAGTTCCTCTTACTGGGATTTCC 50

17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
   |||||||||||||||||||||||||||||||||||||||||||||||||||
51 CGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTTCTTCTCCCTGTTCT 100

34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGluIleSerLeu 50
   |||||||||||||||||||||||||||||||||||||||||||||||
101 ACGTCTTCACCCTGCTTGGGGAACGGGACCATACTGGGCGTCTCACTCTCACTG 150

51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
   ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GACTCCAGACGTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCGGT 200

67 lValAspIleAlaTyrAlaAlaCysAsnThrValProArgMetLeuValAsnL 84
   |||||||||||||||||||||||||||||||||||||||||||||||||||||
201 CGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACC 250

84 euLeuHisProAlaLysProIleSerPheAlaGlyAlaArgMetMetGlnThr 100
   ||||||||||||||||||||||||||||||||||||||||||||||||||||
251 TCCTGCATCCACCCAGCCAAGCCCATCTCCTTTGCGGGCGCCAGAATGATGCAGACC 300

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValVal
   ||||||||||||||||||||||||||||||||||||||||||||||||| Me 117
301 TTTCTGTTTTCCACTTTTGCTGTCACAGAAGTGCTCCTCCTGGTGGTCAT 350

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 GTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA 400

134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
    |||||||||||||||||||||||||||||||||||||||||||||||||
401 TCATGACCTGGAGAGTGTGCATCACCCTCGCGGTGACTTCCTGGACCACT 450

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167
    |||||||||||||||||||||||||||||||||||||||||||||||
451 GGAGTCCTTTTATCCCTTGATTCATCTGTGTTACTTCCTTTACCCTT 500

167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
    |||||||||||||||||||||||||||||||||||||||||||||||||
501 CTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAAAACTTGGCTGTTC 550

---

184 euLysLeuUalCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 TCAAACTTGCCTGTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 600

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 GGAGCAAATTCTGGGCTGGTGGGACCCTTGTCCACAATTGTAGTTCATA 650

217 rMetCysIleLeuLeuCysSerAlaIleLeuGlnIleSerArgGluValGln 234
    |||||||||||||||||||||||||||||||||||||||||||||||||||
651 TATGTGCATCCTCTGTGCTATCCTCAGAATCCAATCAAGGGAAGTTGAA 700

234 rGlyLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
    |||||||||||||||||||||||||||||||||||||||||||||||||||
701 GGAAAGCCTTCCGCACCTGCTTCTCCCACCTGTGTGTTGATTGGACTCGTT 750

251 TyrGlyThrAlaIleIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 TATGGCACAGCCATTATCATCATGTATGTTGGACCCAGATATGGGAACCCCAA 800

267 sGluGlnLeuArgGluLeuThrArgAlaLeuThrSerThrTyr 283
    |||||||||||||||||||||||||||||||||||||||||||
801 GGAGCAGAAGAAATATCTCCGCTGTTTCACAGCTCTTAAATCCATGC 850

284 euAsnProLeuLeuCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 TCAATCCCCTTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTG 900

301 LysArgValLeuGlyValGlyUArgAlaLeu 310
    ||||||||||||||||||||||||||||||
901 AAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 930

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:AAS09946

seq_documentation_block:
ID   AAS09946 standard; DNA; 1040 BP.
XX
AC   AAS09946;
XX
DT   24-OCT-2001 (first entry)
XX
DE   DNA encoding human odorant receptor (OR)-like protein, NOV2.
XX
KW   Odorant receptor; human; OR; NOV2; therapeutic; neuro-olfactory system;
KW   trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;
KW   lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;
KW   acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;
KW   Albright hereditary osteodystrophy; diagnostic; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   5'UTR   1..81
FT   CDS   82..1014
FT   /*tag= a
FT   /*tag= b
FT   /product= "Odorant receptor-like protein, NOV2"
FT   3'UTR   1015..1040
FT   /*tag= c
XX
PN   WO200151632-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01513.
XX
PR   13-JAN-2000; 2000US-0175989.
PR   14-JAN-2000; 2000US-0176134.
PR   25-JAN-2000; 2000US-0177839.
PR   26-JAN-2000; 2000US-0178191.
PR   26-JAN-2000; 2000US-0178227.
PR   14-JUL-2000; 2000US-0218324.

```
XX
PT  New polynucleotides which encode polypeptides involved in olfactory
PT  sensation for identifying olfactory agonists and antagonists -
XX
PS  Claim 8; Page 251; 1857pp; English.
XX

Q   The present sequence is one of a number of isolated polynucleotides
CC  which encode polypeptides involved in olfactory sensation. The
CC  polynucleotides can be used in screening for olfactory agonists and
CC  antagonists. The methods allow for the determination of primary
CC  scents and the identification of the odour receptors used to detect
CC  these primary scents. The methods also enable determination of
CC  secondary scents and the identification of combinations of odour
CC  receptors that are involved in detecting such secondary scents.
CC  This enables the construction of a scent representation (also called
CC  a scent fingerprint or scent profile), which may be used to re-create
CC  and edit scents. Libraries of olfactory receptors are useful for
CC  determining the interaction pattern of a composition with the receptors,
CC  and can be used for determining differences in the olfactory faculties
CC  of different individuals.
XX

Q   Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

alignment_scores:
    Quality:  1584.00          Length:     310
    Ratio:    5.160            Gaps:       0
    Percent Similarity: 99.032 Percent Identity: 99.032

alignment_block:
US-09-898-586-24 x AAH31617

Align seq 1/1  to: AAH31617    from: 1  to: 930

                AAH31617    from: 1  to: 930       .

1   MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1   ATGGGAGACAATATAACATCCATCACAGAGAGTTCCTCACTGGGGATTTCC 50

17  oValGlyProCysArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
    ||||||||||||||||||||||||||||||||||||||||||||||||||
51  CGTTGGGCCCAAGGAATTCAGATGCTTCTCTTTGGGCTCTTCTCCCTGTTCT 100

34  yrValPheThrLeuLeuGlyAsnGlyThrIleLeuLeuIleLeuSerLeu 50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 ACGTGTTCACCCTGCTGGGGAACGGGACCATCTTGCTCATCCTGTCACTG 150

51  AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCCGT 200

67  lValAspIleAlaThrValAlaCysAsnThrValProArgMetLeuValAsnL 84
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 CGTCGACATCGCCACACGGTGCCCCCGGATGCTGGTGAACC 250

84  euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGAGCAGACC 300

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCCTGGTGGTGAT 350

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 GTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA 400

134 leMetThrArgProArgValCysIleThrLeuAlaValThrSerIleProThr 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 TCATGACCTGGAGAGTCTGCATCACCCTGGCTGTGACTTCCTGGACCACT 450

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProh 167
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 GGAGTCCTTTTATCCTGATCCATCTTGTGTTACTTCTACCTTTACCCCTT 500

167 eCysArgProGlnLysIleThrTyrHisPhePheCysGluIleLeuAlaVaL 184
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 CTGTAGGCCCCAGAAAATTACCTATCACTTTTTTTGTGAAATCTTGGCTGTTC 550

184 euLysLeuAlaCysSerAlaAspThrHisLeuAsnGluAsnMetValLeuAla 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
551 TCAAACTTCGCCTGTCCAGATACCCATCAATGAGAACATGGTCTTGGCC 600

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
    ||||||||||||||||||||||||||||||||||||||||||||||||||
601 GGAGCAAATTTCTGGGCTGGTGGGACCCTTGTCCACAATTGTRAGTTTCATA 650

217 rMetCysIleLeuCysAlaIleIleLeuGlnIleGlnSerArgGluValGlnA 234
    ||||||||||||||||||||||||||||||||||||||||||||||||||
651 TATGTGCANCCTCTGTGCTATCCTTCAGATCCAATCAAAGGGAAGTTTCAGA 700

234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
701 GGAAAGCCTTCCCACCCTGCTTCTCCCACCCTCGTGTGATTGGACTCGTT 750

251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
    ||||||||||||||||||||||||||||||||||||||||||||||||||
751 TATGGCACAGCCATTATCATGTATGTTGGACCCAGAATATGGCAACCCAA 800

267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetL 284
    ||||||||||||||||||||||||||||||||||||||||||||||||||
801 GGAGCAGAAGAAAATATCCTCTGCGTTTCACAGCCCTCTTTAATCCCATGC 850

284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
851 TCAATCCCCTTATCTGTAGTCTTTAGGAACTCAGAAGTGAAGAATACTTTG 900

301 LysArgValLeuGlyValGlnArgGlnLeu 310
    ||||||||||||||||||||||||||||||||
901 AAGAGAGTGCTGGGAGTAGAAAGGCTTTA 930
```

```
XX
PI    Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR    WPI; 2001-290713/30.
XX
PT    New polynucleotides which encode polypeptides involved in olfactory
PT    sensation for identifying olfactory agonists and antagonists -
XX
PS    Claim 8; Page 554; 1857pp; English.
XX
CC    The present sequence is one of a number of isolated polynucleotides
CC    which encode polypeptides involved in olfactory sensation. The
CC    polynucleotides can be used in screening for olfactory agonists and
CC    antagonists. The methods allow for the determination of primary
CC    scents and the identification of the odour receptors used to detect
CC    these primary scents. The methods also enable determination of
CC    secondary scents and the identification of combinations of odour
●     receptors that are involved in detecting such secondary scents.
CC    This enables the construction of a scent representation (also called
CC    a scent fingerprint or scent profile), which may be used to re-create
CC    and edit scents. Libraries of olfactory receptors are useful for
CC    determining the interaction pattern of a composition with the receptors,
CC    and can be used for determining differences in the olfactory faculties
CC    of different individuals.
XX
SQ    Sequence 930 BP; 192 A; 266 C; 201 G; 271 T; 0 other;

alignment_scores:
                              Length:        310
            Quality: 1602.00  Gaps:          0
            Ratio:   5.184
Percent Similarity: 99.677    Percent Identity: 99.677

alignment_block:
US-09-898-586-24 x AAH32366

Align seg 1/1  to: AAH32366  from: 1  to: 930

                              AAH32366  from: 1  to: 930

1    MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr  17
     |||||||||||||||||||||||||||||||||||||||||||||||
1    ATGGGGGACAATATAACATCCATCACACAGAGTTCCCCTACTGGGATTTCC  50

17   oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT  34
     ||||||||||||||||||||||||||||||||||||||||||||||
51   CGTTGGCCCAAGGAATTCAGAAGCTCCTCTTTGGGCTTCTTCTCCGTTTCT  100

34   yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu  50
     ||||||||||||||||||||||||||||||||||||||||||||||
101  ACGTCTTCACCCTGCTCGGGGAACGGGACCATACTGGGGCTCATCTCACTG  150

51   AspSerArgLeuHisAlaProMetThrTyrPhePheLeuSerHisLeuAla  67
     ||||||||||||||||||||||||||||||||||||||||||||||
151  GACTCCAGAGTCCACGCCCCCATGACTTACTTCTTCCTCTCACACCTGGCGT  200

67   ValAspIleAlaTrpIleAlaCysAsnThrValProLeuMetLeuValAsnL  84
     ||||||||||||||||||||||||||||||||||||||||||||||
201  CGTTGGACAATCGCCTACGGGCTGCCCAACACGGTGCCCCGGATGCTGGTGAACC  250

84   euLeuHisProAlaLysProIleSerPheAlaGlyTyrArgMetMetGlnThr  100
     ||||||||||||||||||||||||||||||||||||||||||||||
251  TCCTGCATCCAGCCAAGCCAATCTCCTTTGCGGGCCGCATGATGCAGACC  300

101  PheLeuPheSerThrPheAlaValThrGluLeuSerLeuLeuValValMe  117
     ||||||||||||||||||||||||||||||||||||||||||||||
301  TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCCTCCTGGTTGTGAT  350

117  tSerTyrArgAspLeuTyrValAlaIleLeuCysHisProLeuArgTyrLeuAlaI  134
     ||||||||||||||||||||||||||||||||||||||||||||||
351  GTCCTATAGATCTTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA  400

134  leMetThrTrpArgValCysIleThrLeuAlaValIhrSerTrpPhThrIhr  150
     ||||||||||||||||||||||||||||||||||||||||||||||
401  TCATGACCTGGAGAGTCTGCATCACCCTCGCGGTGACTTCCTGACCACT  450
```

```
151  GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh  167
     ||||||||||||||||||||||||||||||||||||||||||||||
451  GGAGTCCTTTTACCCTGATTCATCTTGTGTTACTTCTACCTTTACCCT  500

167  eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVaIL  184
     ||||||||||||||||||||||||||||||||||||||||||||||
501  CTGTAGGCCCCAGAAAAATTTATCACTTTTTTTGTGAAAICTTGGCCTGTTC  550

184  eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla  200
     ||||||||||||||||||||||||||||||||||||||||||||||
551  TCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC  600

201  GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy  217
     ||||||||||||||||||||||||||||||||||||||||||||||
601  GGAGCAATTTCTGGGCTGGTGGGACCCTTGTCCACAAATGTAGTTGTCATA  650

217  rMetCysIleLeuCysAlaIleLeuLeuGlnIleLeuSerArgGlnValGln  234
     ||||||||||||||||||||||||||||||||||||||||||||||
651  TATGTGCATCCTCGTCGTTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA  700

234  rglysAlaProPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe  250
     ||||||||||||||||||||||||||||||||||||||||||||||
701  GGAAAGGCTTTCGCCACCTGCTTCTCCCACCTCTGCGTGATTGGACTCTTT  750

251  TyrGlyThrAlaIleIleMetThrTyrValGlyProArgTyrGlyAsnProLy  267
     ||||||||||||||||||||||||||||||||||||||||||||||
751  TATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAA  800

267  sGluGlnGlyLysTyrLeuLeuPheHisSerLeuPheAsnProMetL  284
     ||||||||||||||||||||||||||||||||||||||||||||||
801  GGAGCAGGGAAAATATCTCCCGCTGTTTCACAGCCTCTTTAATCCCATGC  850

284  euAsnProLeuIleCysSerLeuArgAsnSerGluValLysSerAsnThrLeu  300
     ||||||||||||||||||||||||||||||||||||||||||||||
851  TCAATCCCCTTTATCTGTAGTCTTAGGAACTCAGAGAGTGAAGAATACTTTG  900

301  LysArgValLeuGlyValGluArgAlaLeu  310
     ||||||||||||||||||||||||
901  AAGAGAGTGCTGGGAGTAGAAAGGGCTTTA  930
```

```
seq_name: /SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH31617
ID    AAH31617 standard; DNA; 930 BP.
XX
AC    AAH31617;
XX
DT    30-JUL-2001  (first entry)
XX
DE    Human olfactory receptor polynucleotide, SEQ ID NO: 190.
XX
KW    Human; olfactory receptor; OR; primary scent determination;
KW    secondary scent determination; polypeptide library; odour receptor;
KW    scent profile; scent fingerprint; scent representation; ds.
XX
OS    Homo sapiens.
XX
PN    WO200127158-A2.
XX
PD    19-APR-2001.
XX
PF    06-OCT-2000; 2000WO-US27582.
XX
PR    08-OCT-1999; 99US-0158615.
PR    24-FEB-2000; 2000US-0184809.
XX
PA    (DIGI-) DIGISCENTS.
PA    (YEDA ) YEDA RES & DEV CO LTD.
XX
PI    Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR    WPI; 2001-290713/30.
```

```
PI  Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;
PI  Spytek KA, Li L;
XX
DR  WPI; 2001-451859/48.
DR  P-PSDB; AAU05142.
XX
PT  New NOVX polypeptides and polynucleotides, useful for treating or
PT  preventing disorders of the neuro-olfactory system, cancer and multiple
PT  sclerosis.
XX
PS  Claim 9; Page 55; 141pp; English.
XX
CC  The sequence represents the coding sequence of human odorant receptor
CC  (OR)-like protein, NOV12. The NOV12 polypeptide, nucleic acid and
CC  antibody are useful as therapeutics, particularly in the manufacture of a
CC  medicament for treating a syndrome associated with a human disease, which
CC  includes a pathology associated with NOV12 polypeptide. The NOV12 nucleic
CC  acid and polypeptide are especially useful in therapeutic or
CC  prophylactic applications for disorders of the neuro-olfactory system,
CC  e.g. those induced by trauma, surgery and/or neoplastic disorders. The
CC  DNA encoding the protein is useful in gene therapy for treating the
CC  above conditions. Furthermore, the nucleic acids and polypeptides are
CC  useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus
CC  cancer, immune response, acquired immunodeficiency syndrome (AIDS),
CC  asthma, Crohn's disease, multiple sclerosis or Albright hereditary
CC  osteodystrophy. These are also useful in developing powerful assay
CC  system for functional analysis of various human disorders, as well as
CC  in diagnostic applications.
XX
SQ  Sequence 1014 BP; 219 A; 281 C; 216 G; 298 T; 0 other;
```

```
alignment_scores:
  Quality:      1609.00        Length:         310
  Ratio:           5.190       Gaps:             0
Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
US-09-898-586-24 x AAS09956

Align seq 1/1  to: AAS09956   from: 1  to: 1014

AAS09956 ..
```

```
   1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr  17
     |||||||||||||||||||||||||||||||||||||||||||||||||||
   1 ATGGGGGACAATAATAACATCACAAGTATCACACAGAGTTCCTCCTACTGGGGATTTCC 104

  17 oValGlyProArgIleGlnIleMetLeuPheGlyLeuPheSerLeuPheT  34
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 105 CGTTGGGCCAAGGATTCAGATGCTCTCTTTTGGGCTCTTCTCCCGTTTCT 154

  34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyIleLeuSerLeu  50
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 155 ACGTCTTCACCCTGCTGGGAAACGGGACCATACTGGGGATCTTATCACTG 204

  51 AspSerArgLeuHisAlaProMetTyrPheLeuLeuSerLeuAla  67
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 205 GACTCCAGACTCCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCCGT 254

  67 lValAspIleAlaTyrValAlaCysAsnThrValProArgMetLeuValAsnL  84
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 255 CGTCGACATCGCCTACGCGTGCAACACGGTGCCCCGGATGCTGGTGAACC 304

  84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 305 TCCTGCACCCAGCCAAGCCCATCTCCTTCGCGGGCCGCATGATGCAGACC 354

 101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValAlaMe 117
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 355 TTTCTGTTTTCCACTTTTGCTGTCACAGAATGTCTCCTCCTGGTGGCTAT 404

 117 tSerThrTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla I 134
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 405 GTCCTATGATCTGTACGTGGCCATCTGCCACCCCCTCCGATATTTGGCCA 454
```

```
seq_name: /SIDS1/gcgdata/hold-geneseq-embl/NA2001A.DAT:AAH32366

seq_documentation_block:
ID  AAH32366 standard; DNA; 930 BP.
XX
AC  AAH32366;
XX
DT  30-JUL-2001  (first entry)
XX
DE  Human olfactory receptor polynucleotide, SEQ ID NO: 939.
KW  Human; olfactory receptor; OR; primary scent determination;
KW  secondary scent determination; polypeptide library; odour receptor;
KW  scent profile; scent fingerprint; scent representation; ds.
OS  Homo sapiens.
XX
PN  WO200127158-A2.
XX
PD  19-APR-2001.
XX
PF  06-OCT-2000; 2000WO-US27582.
XX
PR  08-OCT-1999;  99US-0158615.
PR  24-FEB-2000; 2000US-0184809.
XX
PA  (DIGI-) DIGIGENTS.
PA  (YEDA ) YEDA RES & DEV CO LTD.
```

```
 134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 455 TCATGACCTGGAGAGTCTGCATCACCCTCGCCGTGACTTCCTGGACCACT 504

 151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProHi 167
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 505 GGAGTCCTTTTTATCCTTGCATTCACCTGTGTTACTTCTACCTTTACCCTT 554

 167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 555 CTGTAGGCCCCCAGAAAATTTATCACTTTTTTGTGAAATCTTGGCTGTTC 604

 184 euLysLeuValAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 605 TCAAACTTGGCTGCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 654

 201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 655 GGAGCAAATTCTGGGCTGGTGGGACCCTTGTCCACAATTGTTGTTTCATA 704

 217 rMetCysIleLeuVAlAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 705 TATGTTGCATCCTCTGTGCTATCCTTCAGATCCAATCCAAGGGAAGTTCAGA 754

 234 rgLysAlaPheCysThrCysPhePheSerHisLeuCysValIleGlyLeuPhe 250
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 755 GGAAAGGCTTTCTGCACCTGCTTCTTCCACCTGTGTGTGATAGGACTCTTT 804

 251 TyrGlyThrAlaIleIleLeuMetThrTyrValGlyProArgTyrGlyAsnProLy 267
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 805 TATGGCACAGCCATTATCCATGTAATGTTGGACCCAGATAGGGAAACCCCAA 854

 267 sGlnGlnLysLysTyrLeuAlaLeuPheHisSerLeuPheAsnProMetL 284
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 855 GGACCAGAAGAAAATATTCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGC 904

 284 euAsnProLeuIleCysSerLeuArgAsnSerGluVallLysAsnThrLeu 300
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 905 TCAATCCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGAATACTTTG 954

 301 LysArgValValGluGlyValGluArgAlaLeu 310
     |||||||||||||||||||||||||||||||||
 955 AAGAGAGTGCTGGGAGTAGAAAGGGCTTTA 984
```

```
CC   acid and polypeptide are especially useful in therapeutic or
CC   prophylactic applications for disorders of the neuro-olfactory system,
CC   e.g. those induced by trauma, surgery and/or neoplastic disorders. The
CC   DNA encoding the protein is useful in gene therapy for treating the
CC   above conditions. Furthermore, the nucleic acids and polypeptides are
CC   useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus
CC   cancer, immune response, acquired immunodeficiency syndrome (AIDS),
CC   asthma, Crohn's disease, multiple sclerosis or Albright hereditary
CC   osteodystrophy. These are also useful in developing powerful assay
CC   system for functional analysis of various human disorders, as well as
CC   in diagnostic applications.
XX
SQ   Sequence 1012 BP; 219 A; 281 C; 215 G; 297 T; 0 other;


alignment_scores:
  Quality:    1609.00                 Length:   310
  Ratio:         5.190                Gaps:       0
  Percent Similarity: 100.000     Percent Identity: 100.000

alignment_block:
  US-09-898-586-24 x AAS09955

Align seg 1/1 to: AAS09955   from: 1 to: 1012


   1 MetGlyAspAsnIleThrSerIleThrGluProLeuLeuLeuGlyPhePr 17
     ||||||||||||||||||||||||||||||||||||||||||||||
  54 ATGGGAGACAATATAACATCCATCACAGAGCTTCCTCACTGGGATTTCC 103

  17 oValGlyProArgIleGlnMetLeuLeuLeuPheGlyLeuLeuPheSerLeuPheThr 34
     |||||||||||||||||||||||||||||||||||||||||||||
 104 CGTTGGCCCAGGAATTCAGATGCTCCTCCTTTGGGCTTCTTTCCCCTGTTCT 153

  34 rValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
     |||||||||||||||||||||||||||||||||||||||||
 154 ACGTCTTCACCCTGCTCGGGGAACGGGACCATACTGGGGCTCATCTCACTG 203

  51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
     ||||||||||||||||||||||||||||||||||||||||||
 204 GACTCCAGACTGCACGCCCCCATGTACTTCTTCCTCTCACACCTGGCGGT 253

  67 lValAspIleLeuAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
     |||||||||||||||||||||||||||||||||||||||||
 254 CGTCGACATCCTGGCTTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACC 303

  84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr 100
     ||||||||||||||||||||||||||||||||||||||
 304 TCCTGCATCCAGCCAAGCCCATCTCCTTTGCGGGCCGCATGAGTCCAGACC 353

 101 PheLeuPheSerThrPheAlaValThrGluCysGlyLeuLeuLeuValAlaMe 117
     |||||||||||||||||||||||||||||||||||||||||
 354 TTTCTGTTTTCCACTTTTGCTGTCACAGAAGTCTCCCTCGGTGGTGAT 403

 117 tSerTyrAspLeuPheTyrValAlaIleCysHisProLeuArgTyrLeuAlaIl 134
     ||||||||||||||||||||||||||||||||||||||||
 404 GTCCTATGACTTGTACGTGGCCCATCCACCCCTCCGGATATTTGGCCA 453

 134 leMetThrThrProArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
     |||||||||||||||||||||||||||||||||||||||
 454 TCATGACCTGGAGAGTCGATCACCCTGGCGGTGACTTCCTGGACCACT 503

 151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProh 167
     |||||||||||||||||||||||||||||||||||
 504 GGAGTCCTTTTATCCTTGATTCATCTTGTGTTACTTCCACTTTACCCTT 553

 167 eCysArgProGlnLysIleTyrHisPheCysGluIleLeuAlaValL 184
     |||||||||||||||||||||||||||||||
 554 CTGTGAGGCCCCAGAAAAATTTATCACTTTTTTGTGAAATCTTGGCTGTTC 603

 184 euLysLeuAlaCysSerAspThrHisIleAsnGluAsnMetValLeuAla 200
     |||||||||||||||||||||||||||||||||||||||||
 604 TCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 653

 201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrLeuValValSerTy 217
     |||||||||||||||||||||||||||||||||||||||||
 654 GGAGCAAATTCTGGCGTCGGTGGGACCCTGTTCCACAATTGTAGTTGCATA 703

 217 rMetCysIleLeuCysValAlaIleLeuGluIleLeuGlnSerArgGluValGlnA 234
     |||||||||||||||||||||||||||||||||||||||||||||
 704 TATGTCGCATCCTCTGTGCTAACCTTCAGATCCAATCAAGGGAAGTTCAGA 753

 234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPheLeuPhe 250
     ||||||||||||||||||||||||||||||||||||||||||||
 754 GGAAAGCCTTTCTGCACCTGCTTCTCCCACCTCTGCGTTATTGGACTGTTTT 803

 251 TyrGlyThrAlaIleIleMetThrTyrValGlyProArgTyrGlyValAsnProLy 267
     ||||||||||||||||||||||||||||||||||||||||||||
 804 TATGGCACAGCCATTATCATGACATATGTTGGACCCAGAATATGGGAACCCAA 853

 267 sGluGlnLysLysTyrTyrLeuLeuLeuPheHisSerSerLeuPheAsnProMetL 284
     |||||||||||||||||||||||||||||||||||||||||||||
 854 GGAGCAGAAGAAAATATCTCCTGCTGTTTCACAGCTTCTTTAATCCCATGC 903

 284 euAsnProLeuIleCysSerLeuArgAsnGluGluValLysAsnThrLeu 300
     ||||||||||||||||||||||||||||||||||||||||
 904 TCAATCCCCTTATCTGTTAGTGTCTTAGGAACTCAGAAGTGAAGAATACTTTG 953

 301 LysArgValLeuGlyValGlyValArgAlaLeu 310
     ||||||||||||||||||||||||||||
 954 AAGAGAGTGCTCGGGAGTAGAAAGGGCTTTA 983
```

```
                            Ratio:                    Gaps:
Percent Similarity: 100.000   Percent Identity: 100.000   Gaps: 0   Ratio: 5.190

alignment.block:
US-09-898-586-24 x AAF58613

Align seg 1/1 to: AAF58613 from: 1 to: 974

   1 MetGlyAspAsnIleThrSerIleThrGlyPheLeuLeuLeuGlyPhePr  17
     ||||||||||||||||||||||||||||||||||||||||||||||||||
   1 ATGGGGGACAANATAACATCCATCACAGGATTTCTTCCTACTGGGATTTCC  50

  17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuLeuGlyPhePheT  34
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  51 CGTTGGCCCAAGGAATTCAGAGATGCTTCTTCTTTGGGCTTCTTCCGTTCT 100

  34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuIleSerIleLeu     50
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 101 ACGTCTTCACCCTGCTCGGGAACGGGACCATACTGATCTCAATCCTGTCTG 150

  51 AspSerArgLeuHisAlaProMetThrTyrPhePheLeuSerHisLeuAlaVa 67
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 151 GACTCCAGACTGCACGCCCCATGACTTACTTCTTCTCTCACACCTGGGCGGT 200

  67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn.. 84
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 201 CGTCGACATCGCCTACGCCTGCAACACGGTGCCCCGGATGCTGGTGAACC  250

  84 euLeuHisProAlaLysProIleSerPheProAlaGlyArgMetMetGlnThr 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 TCCTGCAYCCAGCCAAGCCCATCTCCTTTCCGGCGGGCCGCATGATGCAGACC 300

 101 PheLeuPheSerThrPheAlaValThrHisGlyCysLeuLeuLeuLeuValMe 117
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 TTTCTGTTTTCCACTTTTGCTGTCACACATGGCTGCCTCCTGCTGGTGAT  350

 117 tSerThrTyrAspLeuHisTyrValAlaIleLeuAlaIle            134
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 351 GTCCTCTATGATCTGCATTACGTGGCCATCCTGGCCATCCCGTATTGGCCA 400

 134 leMetThrTrpArgValCysIleThrLeuLeuAlaValThrSerTrpThrThr 150
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 401 TCAATGACCTGGAGAGTGTGCANCACCCTGCTGGCTGTGACTTCCTGGACCACT 450

 151 GlyValLeuSerGlyLeuValGlyLeuProLeuSerThrPheAlaValThrP 167
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 451 GGAGTCCTTTTATCCTTGATTCATCTGTGTCTTGTTCTTCTACCTTTACCCTT 500

 167 eCysArgProGlnLysIleIleThrPheLeuSerLysSerTrpLeuVall  184
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 501 CTGTAGGCCCCAGAAAATTTATCACTTTTTTGTCAAAATCTTGGCTGGTTC 550

 184 euLysLeuAlaCysAlaAspThrHisIleSerIleGluAsnGluAsnMetValLeuGlyA 201
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 551 TCAAACTTGCCTGTCGAGATACCCACATCAATAGAAACATGGGTCTTGGCC 600

 201 GlyAlaIleLeuSerGlyLeuValGlyLeuProLeuSerThrIleValValSerTy 217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 601 GGAGCAATTTCTGGGCTGGGCACCCCTTGTCCACAATTGTAGTTTCATA  650

 217 rMetCysIleLeuCysAlaIleLeuGlnIleLeuSerArgArgGluValGlnIle 234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 651 TACGTGGCATCCCGTGTCCTATCCCTTACCAATCAAGGGAAGTTCAGA  700

 234 rgLysAlaAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 701 GGAAAAGCCTTCTGCACCTGCTTCTCCCACCTCTGTGATTGGACTCTTT  750

 251 TyrGlyThrAlaIleIleMetTyrValGlyLeuAsnProLy           267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 751 TATGGCACAGCCAATTATCATGTATGTTGGACCCAGATATGGGACCCCAA 800

 267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetLl    284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 801 GGACCAGAAGAAAATATCTCCTGCTGTTTCACAGCCTTCTTTAATCCCATGC 850

 284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu  300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 851 TCAATCCCCTTATCTGTGAGTCTTAGGAACTCAGAAGTGAAGAAATACTTTG 900

 301 LysArgValLeuGlyValGluArgGlyAlaLeu                   310
     ||||||||||||||||||||||||||||||||||
 901 AAGAGAGTGCTCTGGGAGTGAGAAAGGGCTTTA                  930
```

seq_name: /SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS09955

```
seq_documentation.block:
ID   AAS09955 standard; DNA; 1012 BP.
XX
AC   AAS09955;
XX
DT   24-OCT-2001 (first entry)
XX
DE   DNA encoding human odorant receptor (OR)-like protein, NOV11.
XX
KW   Odorant receptor; human; OR; NOV11; therapeutic; neuro-olfactory system;
KW   trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;
KW   lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;
KW   acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;
KW   Albright hereditary osteodystrophy; diagnostic; ds.
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FH
FT   5'UTR            1..53
FT                    /tag= a
FT   CDS              54..986
FT                    /tag= b
FT                    /product= "Odorant receptor-like protein, NOV11"
FT   3'UTR            987..1012
FT                    /tag= c
XX
PN   WO200151632-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01513.
XX
PR   13-JAN-2000; 200US-0175989.
PR   14-JAN-2000; 200US-0176134.
PR   25-JAN-2000; 200US-0177839.
PR   26-JAN-2000; 200US-0178191.
PR   26-JAN-2000; 200US-0178227.
PR   14-JUL-2000; 200US-0218324.
PR   24-JUL-2000; 200US-0220253.
PR   25-JUL-2000; 200US-0220590.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;
PI   Spytek KA, Li L;
XX
DR   WPI; 2001-451859/48.
DR   P-PSDB; AAU05141.
XX
PT   New NOVX polypeptides and polynucleotides, useful for treating or
PT   preventing disorders of the neuro-olfactory system, cancer and multiple
PT   sclerosis.
XX
PS   Claim 9; Page 52; 141pp; English.
XX
CC   The sequence represents the coding sequence of human odorant receptor
CC   (OR)-like protein, NOV11. The NOV11 polypeptide, nucleic acid and
CC   antibody are useful as therapeutics, particularly in the manufacture of a
CC   medicament for treating a syndrome associated with a human disease, which
CC   includes a pathology associated with NOV11 polypeptide. The NOV11 nucleic
```

OM of: US-09-898-586-24 to: N_Geneseq_032802:*   out_format : pfs

Date: Aug 21, 2002   8:22 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL-frame+_p2n.model -DEV-xlh
-O-/cgn2_1/USPTO_spool/US09898586/runat_16082002_205736_20603/app_query.fasta_1.373
-DB-N_Geneseq_032802 -QFMT-fastap -SUFFIX-rng -GAPOP=12.000
-GAPEXT-4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPEXT=4.500 -QGAPEXT-0.050 -XGAPEXT-0.500
-QGAPOP=6.000 -FGAPEXT-7.000 -XGAPOP-10.000 -YGAPEXT-0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-DELEXT=7.000 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pts
-TRANS-human40.cdi -LIST-45 -DOCALIGN-200 -THR_SCORE=pct
-NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN 2000000000
-USER=US09898586.@CGN1_1.169 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXX -WAIT -THREADS=1

Search information block:
Query: US-09-898-586-24
Query length: 310
Database: N_Geneseq_032802:*
Database length: 85457221
Database sequences: 1736436
Search time (sec): 199.760000

score_list:

Sequence	Strd Orig	Zscore	EScore	Len t Documentation
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAF58613	+	1609.00	2938.22	2.6e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH31617	+	1609.00	2937.82	2.7e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS09955	+	1609.00	2937.80	2.7e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS09956	+	1602.00	2925.84	1.3e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32366	+	1584.00	2892.75	8.8e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAB31648	+	1584.00	2892.31	8.8e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS09946	+	1584.00	2891.57	1.0e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS09343	+	1583.00	2887.26	1.8e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS08660	+	1147.00	2079.93	1.7e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAB09073	+	1147.00	2079.93	1.7e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS08655	+	1138.00	2065.96	1.0e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS08654	+	1138.00	2061.47	1.0e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32363	+	1126.00	2050.87	7.0e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32305	+	1126.00	2050.84	7.0e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAF58609	+	1126.00	2050.80	7.0e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS08657	+	1080.00	1959.37	8.6e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA1996.DAT:AAT33900	+	1048.50	1901.88	8.6e-1
/SIDS1/gcgdata/hold-geneseqn-embl/NA2000.DAT:AAA70768	+	1048.50	1901.88	1.4e-97
/SIDS1/gcgdata/hold-geneseqn-embl/NA2000.DAT:AAZ60116	+	1048.50	1901.88	1.4e-97
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32170	+	992.00	1804.33	3.7e-9
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS42450	+	991.00	1802.59	4.7e-9
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS42444	+	947.50	1722.66	1.3e-8
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32061	+	946.00	1719.26	2.0e-8
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS42451	+	930.00	1690.45	8.2e-8
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32345	+	920.00	1672.03	8.7e-8
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS42449	+	881.00	1600.34	8.6e-8
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32210	+	881.00	1600.00	8.7e-8
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32119	+	854.00	1555.83	2.6e-7
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32020	+	852.50	1547.79	2.6e-7
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS42045	+	813.00	1475.24	4.7e-7
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32119	+	698.00	1263.93	7.3e-7
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAS42428	+	693.50	1255.17	1.4e-6
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAF58613	+	692.00	1252.87	1.9e-6
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32244	+	690.50	1253.97	1.7
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH31722	+	680.50	1231.66	2.9
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH31846	+	677.00	1225.46	6.5
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH31669	+	672.00	1214.71	2.6
/SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAH32250	+	672.00	1214.71	2.6

seq_name: /SIDS1/gcgdata/hold-geneseqn-embl/NA2001A.DAT:AAF58613

seq_documentation_block:
```
ID   AAF58613 standard; cDNA; 974 BP.
XX
AC   AAF58613;
XX
DT   24-APR-2001 (first entry)
XX
DE   Human RECAP polynucleotide, SEQ ID NO: 41.
XX
KW   Human; RECAP; receptors and associated proteins; cerebroprotective;
KW   nootropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW   antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW   antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW   cytostatic; antibacterial; virucide; fungicide; protozoacide;
KW   antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.
XX
OS   Homo sapiens.
XX
XX
PN   WO200107612-A2.
XX
PD   01-FEB-2001.
XX
PF   21-JUL-2000; 2000WO-US20035.
XX
PR   21-JUL-1999; 99US-0142232.
PR   07-OCT-1999; 99US-0158578.
PR   12-NOV-1999; 99US-0165192.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
PI   Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;
XX
DR   WPI; 2001-168554/17.
DR   P-PSDB; AAB58889.
XX
PT   Novel receptors and associated proteins for diagnosis and treatment of
PT   neurological disorders, immunological disorders including autoimmune/
PT   inflammatory disorders and cell proliferative disorders such as cancer
PT   -
XX
PS   Claim 5; Page 125-126; 128pp; English.
XX
CC   The present sequence encodes a human RECAP (receptors and associated
CC   proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
CC   in the diagnosis, treatment and prevention of neurological disorders
CC   such as stroke, Alzheimer's disease, Pick's disease, Huntington's
CC   disease, dementia, Parkinson's disease, Downs's syndrome, amyotrophic
CC   lateral sclerosis, multiple sclerosis, bacterial and viral meningitis,
CC   CUD (Creutzfeldt-Jakob disease), GSS (Gerstmann-Strassler-Scheinker
CC   syndrome); immunological disorders, including autoimmune/inflammatory
CC   disorders such as AIDS, DiGeorge's syndrome, severe combined
CC   immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
CC   disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
CC   diabetes mellitus, Good Pasture's syndrome, gout, Grave's diseases,
CC   Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
CC   bacterial, fungal, parasitic, protozoal, and helminthic infections; and
CC   cell proliferation disorders such as arteriosclerosis, atherosclerosis,
CC   cirrhosis, hepatitis and cancer.
XX
SQ   Sequence 974 BP; 206 A; 270 C; 214 G; 284 T; 0 other;
XX
```

alignment_scores:
Quality: 1609.00   Length: 310

```
630 TTTCTCTGTTTTGAGTTTTGGACACAGCGAATGTCTTCCTCTGGTGGCTGAT
    |||||       :::||:::    |||||||||||||||||||||||::||
580 GTCCTACGAATCGTTTACGTGGCCAATCTCCCACCCTCTCCGAATACTCCGTCA 531

117 tSerTyrAspLeuThrValAlaIleCysHisProLeuArgThrLeuAlai 134
    |:::||||||||||||||||||||||||||||||||||||:::|||
580 GTCCTACGAATCGTTTACGTGGCCAATCTCCCACCCTCTCCGAATACTCCGTCA 531

134 leMetThrTrpArgValCysIleThrLeuAlaAlaValThrSerTrpThrThr 150
    |||||||||||||||||:::|||||||||||||||||:::||||
530 TCATGACCTGGAGAGTCTGCAATCACCCCTGGGCCTCACTTCCTGACGTGT 481

151 GlyValLeuLeuSerLeuIleHisIleLeuValLeuLeuProLeuProPh 167
    |||       ||||||||||||||||||||||||||||:::||||
480 GGCTCCCCTCCTGGCTCTGGCCCCATGTGGTTCTCATCCTAAGAACTGCCCTT 431

167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVall 184
    |       |||||||||||||||||||||||||||||||::|||
430 CTCTGGGCCCTCATGAAAATCAACACACTTCTTCTGTGAAAATCCTGTGTGTCC 381

184 euLysLeuAlaCysAlaAlaAspThrHisIleLeuAsnGluAsnMetValLeuAla 200
    |::||||||||||||||||||||||||||||       :::|||
380 TCAGGCTGGCCTGTGCTCACCTGGCTCAACCAGGTGGTCATCTTTGCA 331

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
    |||||||||||||||||||       |||       :::|||
330 GCCTGGCTGTTCTTCCTGGGGGCCACCAGCCTGGTGCTTGTCTCCTA 281

217 rMetCysIleLeuCysValAlaIleGlyIleSerArgGluValGlnA 234
    |       :::||||||||||       |||       :::|||
280 CTCGGACATCCTGGCGCGGCCCATCCTGAGGATCCAGTCTGGGGAGGGCGGCA 231

234 rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
    |||||||||||       |||||||||::::||||
230 GAAAAGGGCCTTCTCCACCTGCTTCTCCCACCCTCTGCGGTGGGCACTCTTC 181

251 TyrGlyThrAlaIleIleMetTyrvalGlyProArgTyrGlyAsnProLy 267
    :::||||||||||||       ||||||::::|||
180 TTTGGCAGTGCCATCATCATGTACATGGCCCCCAGTCCCGCCATCCTGA 131

267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
    :||||||||||||||||||||||::::|||
130 GGAGCAGCAAAAGGTCTTTTTTCTATTTTACAGTTTTTTTCAACCCAAAAC 81

284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
    :::||||||||||||||||||||       |||
80 TTAACCCCCTGATTTACAGCCCTGAGGAAACGGAGAGGTCAAGGGTCCCCTG 31

301 LysArgValLeuGlyValGlu 307
    ::::||::||||||     |||
30 AGGAGAGCACTGGGCAAGGAA 10
```